FACULDADE DE ENGENHARIA DA UNIVERSIDADE DO PORTO

Predicting Breast Healing Deformation After Cancer Conservative Treatment

Pedro Miguel Martins de Lemos da Cunha Faria

Mestrado Integrado em Engenharia Informática e Computação

Supervisor: Hélder Filipe Pinto de Oliveira

Co-Supervisor: Hooshiar Zolfagharnasab

September 11, 2017

Predicting Breast Healing Deformation After Cancer Conservative Treatment

Pedro Miguel Martins de Lemos da Cunha Faria

Mestrado Integrado em Engenharia Informática e Computação

September 11, 2017

Abstract

According to the annual report from the World Health Organization, breast cancer is the most frequent cancer among females. Considering all the treatments, surgery is being applied mostly using two methodologies: Mastectomy, that results on removing not only tumor, but also the total breast tissue; and Breast Cancer Surgery (BCS) where only the tumor is removed with a thin layer of healthy tissue around it. It is clear that performing invasive treatment such as surgery, will lead to impose deformations on the breast, which can influence patients' quality of life (QoL). In this way, technology can be assisted to provide a framework that would improve the way patients interact with physicians. Enhancing this framework with the tools to visualize deformation and the healing process after the surgery can elevate patients' QoL.

In order to accomplish the mentioned aim, this thesis focuses on obtaining training models to describe anatomical deformations during the healing process of the breast after BCCT. To achieve reliable training models, a dataset with several 3D breast models is required. Therefore, a semi-synthetic dataset will be generated, containing 3D breast models representing the patients' breasts before and after the surgery. The pre-surgical models are obtained through MRI data of the few patients' data that we have access. The semi-synthetic data of the pre-surgical stage will be generated taking as input these real data and variations of the hypothetic tumor's location and volume and possible breast densities. The pos-surgical data is simulated by a biomechanical wound healing model.

Then by using different machine learning approaches, the relation between the patient's breast before and after the surgery can be obtained and the deformation predicted.

Finally, concerning the evaluation, simulated healed breasts will be compared with the pos-surgical 3D breast models in the dataset through local and global metrics including Euclidean and Hausdorff distances.

Resumo

De acordo com a Organização Mundial de Saúde, o cancro da mama é o cancro mais frequente entre indivíduos do sexo feminino. Tendo em conta todos os tratamentos actualmente disponíveis, a cirurgia é aplicada maioritariamente usando duas metodologias: Mastectomia, que resulta na remoção total da mama e não apenas do tumor; e Tratamento Conservativo do Cancro da Mama no qual apenas é removido o tumor e uma porção reduzida do tecido da mama circundante. Como esperado, a aplicação de tratamento invasivos como o caso da cirurgia leva à deformação da mama afectando a qualidade de vida dos pacientes. Desta forma, a tecnologia poderá ser utilizada de modo a melhorar a interacção entre os pacientes e médicos clínicos de modo a visualizar as possíveis deformações resultantes e o processo de cicatrização após a cirurgia com o intuito de melhorar a qualidade de vida dos pacientes.

De modo a atingir o objectivo acima descrito, é necessário obter modelos de treino capazes de descrever deformações anatómicas ao longo do processo de cicatrização da mama após Tratamento Conservativo do Cancro da Mama. Para que se obtenham modelos de treino viáveis é necessário um dataset com vários modelos 3D. Assim sendo, terá de ser gerado um dataset semi-sintético com modelos 3D representando as mamas das pacientes antes e após a cirurgia. Os modelos pré-cirurgicos serão obtidos com base em informação proveniente de ressonâncias magnéticas das pacientes às quais temos acesso. A informação semi-sintética pré-cirurgica terá em conta a informação real e variações das localizações e volume hipotéticos do tumor e da possível densidade da mama. Os modelos pós-cirurgicos serão simulados com base num modelo biomecânico de cicatrização

Posteriormente, através da utilização de técnicas de aprendizagem computacional, poder-se-á então obter uma relação entre os modelos da mama da paciente antes e após a cirurgia.

Por último, de modo a validar os modelos de previsão, os modelos simulados serão comparados o modelo pós-cirurgico previsto usando diversas métricas como distâncias Euclidianas e de Hausdorff.

Agradecimentos

Em primeiro lugar gostaria de agradecer ao meu orientador, Hélder Oliveira e co-orientador, Hooshiar Zolfagharnasab, pelo apoio e disponibilidade constantes. Ao professor Jaime Cardoso pelas recomendações que foi dando.

Agradeço também à Silvia Bessa e à Sara Oliveira, assim como aos restantes elementos do Breast Research Group, pelos conhecimentos partilhados e pelas sugestões que me foram sendo dadas.

Por último, um especial agradecimento à minha familia e amigos pelo incentivo e compreensão que me têm dado.

Pedro Faria

"At some point, you have to make a decision.
Boundaries don't keep other people out.
They fence you in.
Life is messy.
That's how we're made.
So, you can waste your lives drawing lines.
Or you can live your life crossing them. "

Shonda Rhimes

Contents

List of Figures

List of Tables

Abbreviations

1D	1-dimensional
2D	2-dimensional
3D	3-dimensional
ACR	American College of Radiology
BC	Breast Cancer
BCCT	Breast Cancer Conservative Treatment
BCS	Breast Conserving Surgery
CP	Control Point
CT	Computed Tomography
DCIS	Ductal Carcinoma in Situ
DNN	Deep Neural Networks
FE	Finite Element
FEM	Finite Element Model
FFD	Free Form Deformation
ML	Machine Learning
MLP	Multilayer Perceptron
MOR	Multi-output Regressor
MR	Magnetic Resonance
MRI	Magnetic Resonance Image
LOO	Leave-one-out
QoL	Quality of Life
RF	Random Forest
RFE	Recursive Feature Elimination
RMSE	Root mean-square error

Chapter 1

Introduction

Despite being the most common cancer among females [1], Breast Cancer (BC) is known to be treated with satisfactory aesthetic outcome if diagnosed in early stages. Howsoever, available treatments impose deformation to the treated breasts that impacts the patients' self-esteem and Quality of Life (QoL), that have to live with these consequences for many years. The outcome of the surgery depends on a wide variety of parameters such as the size and location of the tumor, the volume of the breast, the density of both breasts and excised tissue, and finally the effects from complementary treatments such as chemotherapy and radiotherapy and the applied breast cancer surgery.

1.1 Context

Despite of the reasonable aesthetic outcome provided by the current treatments such as BCS (Breast Conserving Surgery), predicting the outcome of the treatment will allow the patient to understand the impacts of treatment and to make a more secure and confident decision. The interaction between physicians and patients would also be enhanced given that the possible aesthetic outcomes and the treatments strategies and the patient's concerns may be discussed through the utilization of visual cues.

1.2 Motivation

There are a few frameworks and tools for the breast plastic surgery planning and even less regarding the oncological surgery planing.

In one side, the existing frameworks on plastic surgery considered transformations that could be applied from female generic torso to high-fidelity 3D models of specific patient. However, this frameworks only focused on breast augmentation and were not designed to preform any other kinds of breast shape deformation.

[1] http://www.wcrf.org/int/cancer-facts-figures/worldwide-data

On the other side, the existing tools on oncological surgery planning used generic female torso that made the surgery planning more difficult once the patient expectations were unreliable, given the greater difficulty of the patients on projecting themselves on the generic torso models, or were based on complex models that required a large computation time being unreliable on real-time scenarios.

The decision making regarding the treatment methodology of BC may be an easier process when it is possible to visualize a predicted deformation of the patient's specific breast model.

1.3 Objectives

This thesis focuses on predicting breast deformation during the healing process after a BCS. In order to allow the prediction of those deformations, machine learning techniques that consider pre-surgery models, simulated post-surgery models and clinical annotations regarding the tumor's and breast's information will be explored. Despite of being already possible to predict these deformations through biomechanical models, this approach will allow to obtain the same predictions on a real-time scenario. To achieve it, and once there is not enough information gathered about real patients, more information is required to be synthesized based on the information of real patients already gathered.

1.4 Contributions

The contributions that this dissertation offers are listed below:

- A semi-synthetic dataset with 3D models of breasts before and after BCS;
- A tool to define the tumor on a patients' 3D model of the breast;
- A model to predict the breast deformations caused by BCS.

1.5 Structure

Besides this chapter, this dissertation counts with five more chapters.

In chapter 2, the fundamentals and the concepts of the breast cancer are explained including: current statistics, most recent and used treatments on this field and the influence and the impact of BCS on the QoL of the patients.

Chapter 3, focuses on 3D models; how the data that would allow theirs generation are gathered, some modeling methodologies techniques and their application with emphasis when used for breast modeling.

Chapter 4, describes the developed tool used in the dataset preparation. This application is described as well as its functional and non-functional requirements.

Chapter 5 presents the applied methodology: generating the dataset, predicting deformations and validating outcomes.

Chapter 6 presents the results obtained trough the thesis development: visual results and metrics for the obtained learning models.

And finally, chapter 7 provides conclusions and possible improvements to enhance the present work in the future.

Chapter 2

Breast Cancer

In this chapter, an overview of the breast anatomy and physiology is presented along side with concepts and treatments regarding breast cancer. The chapter focuses on some of the available types of surgery, and the factors that can influence the aesthetic outcome of those treatments. Not only the anatomical explanations, but also the impacts of breast cancer treatment on the patient's quality of life are described in the following.

2.1 Breast Physiology

The breast is made up of different layers, but predominantly by fat and glandular tissue, having as boundaries the second and the sixth ribs, vertically, and the lateral edge of the sternum and the mid axillary line, medially. Through a woman's aging, the breast suffers transformation during the puberty and menopause, where the ratio between the fat and glandular tissues may vary as well as the blood supply and the lymphatic drainage [EM13]. On adulthood, the mammary gland has about 15 to 20 lobes that converge to the nipple through milk ducts. The milk ducts are surrounded by dense connective tissue, known as fibroglandular tissue or Cooper's ligaments. The majority of the breast, about two thirds, is placed over the pectoral muscle, while its shape is established and maintained by the skin and the Cooper's ligaments [AMS76, VEH+16]. A breast's structure is represented in Figure 2.1.

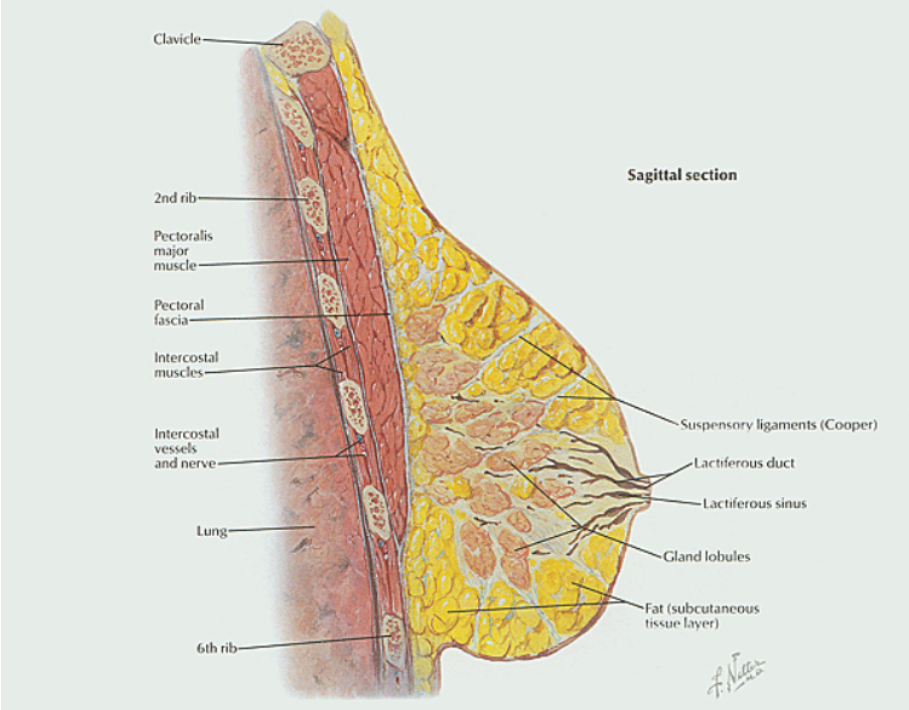


Figure 2.1: Breast's Structure [Wit04]

2.2 Breast Cancer and Incidence

Breast cancer is caused by the presence of a malignant tumor, known as carcinoma, developed from the breast cells. Despite of being more common in women (99%), this disease can also affect men (1%). Portugal follows the same trend as developed countries regarding the diagnosed cases of breast cancer per years. [1]

The tumor, regarding its location, can be associated with one or more of the following 6 regions of the breast: upper-inner, upper-outer, lower-inner, lower-outer, nipple and areola complex or axillary tail. Usually the axillary tail is also considered as upper-outer region. Figure 2.2 shows the breast regions division into quadrants and the nipple and areola complex.

[1] http://www.who.int/cancer/country-profiles/prt_en.pdf?ua=1

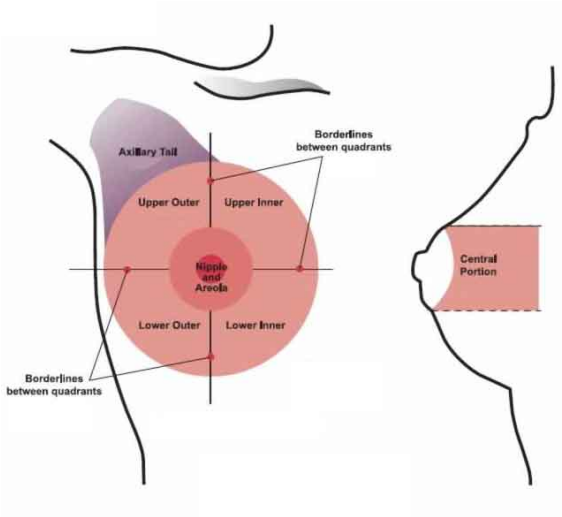


Figure 2.2: Breast regions [2]

According to the study in [SASB08], 57% of the breast cancer occurs in the upper-outer quadrant, 14% in the upper-inner quadrant, 10% in the lower-outer quadrant and 9% in both the lower-inner quadrant and the nipple and areolar complex. The remaining 1% occur in the axilliary tail.

With the increased concerned regarding cancer diseases, it has been made an effort to diagnose them as soon as possible. In the specific case of breast cancer, as shown in Figure 2.3, 85% of the patients who are diagnosed in early stages [3] are granted a 90% chance for the cancer be fully cured and provided with a survival rate greater than 88%. All the diagnosed patients with stage I, II or part of stage III are considered early stages [4]. Early cancer diagnosing allows the patient not only to take less invasive treatments, but also have better aesthetic outcome at the end of treatment.

[2] https://www.pinterest.pt/pin/302304193710434456/

[3] http://www.cancerresearchuk.org/health-professional/cancer-statistics

[4] https://www.womenshealth.gov/publications/our-publications/fact-sheet/early-stage-breast-cancer.html

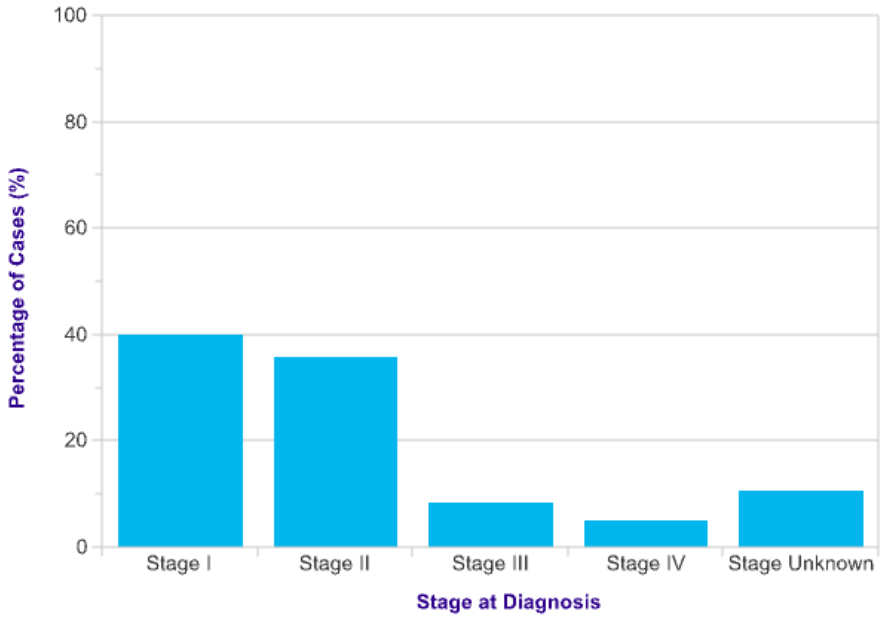


Figure 2.3: Breast cancer incidence by stage [5]

Breast cancer is either considered as both invasive and non-invasive cancer as for spreading to surrounding tissues. The most common example of non-invasive breast cancer is known as Ductal Carcinoma in Situ (DCIS), in this case, the cancer is located on the place where the carcinoma occurred, growing through the milk ducts. If detected in early stages, this type of cancer can be easily cured with a great rate of success, otherwise it can evolve into an invasive breast cancer. On invasive breast cancer, the tumor spreads from the milk ducts and lobules to neighbor tissues. It can be classified in two different types of invasive breast cancer: Invasive Ductal Carcinoma, when the tumor has origin on the milk ducts; or Invasive Lobular Carcinoma, when the tumor has origin on the lobules.

The breast cancer can also be classified in different stages according with the size of the tumor and the affecting tissue. The Stage 0, refers to the DCIS, with a few abnormal cells in lining of the ducts or small portions of the breast, and as a survival rate near 100%; The Stage 1, refers to breast cancer caused by carcinoma with less than 1 inch across (98% survival rate); The Stage 2, refers to breast cancer caused by carcinoma with less than 2 inches across that can spread to some auxiliary lymph nodes (88% survival rate); The Stage 3, refers to breast cancer caused by carcinoma larger than 2 inches across with an extensive spread to auxiliary or nearby lymph nodes. At this stage, some dimpling, inflammation or change of the color skin can be observed (52% survival rate); The Stage 4, refers to breast cancer caused by carcinoma spread from the breast to other regions and organs (16% survival rate); [6]

[5] http://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/breast-cancer

[6] http://johnstonhealth.org/2012/10/breast-cancer-awareness/

One of the most used and successful ways to diagnose breast cancer is thourgh MRI [7]; Mammograms also have an important role evaluating the risk for breast cancer growth, since the medical community has found breast tissue density an important factor for the growth of breast cancer [SSNS12]. Despite of the large number of breast density classifications, a fluently used is the Breast imaging-reporting and Data system (BIRADS) developed by the American College of Radiology (ACR) [SSNS12].

The 4 categories defining the breast's density are enumerated below: [8]

- **1** - fatty: breast is almost entirely fat;
- **2** - scattered fibroglandular: breast has scattered areas of fibroglandular density;
- **3** - heterogeneously dense: breast tissue is heterogeneously dense;
- **4** - dense: breast tissue is extremely dense.

2.3 Breast Cancer Treatment

The goal of breast cancer treatment on early stages of the disease is to completely remove the cancer and preventing its recurrence. On later and more advanced stages of the cancer, it cannot be cured, so, the treatment techniques on this scenario focus on the attenuation of its effects and symptoms together with improving the QoL of the patient.

For years, the mastectomy has been the answer to treat breast cancer. Nowadays there are treatments with better results in terms of tumor removal and with a minor influence on the patients QoL, replacing the mastectomy by treatments such as BCS and oncoplastic treatments.

On the case of breast cancer on men, the mastectomy is always the recommended treatment option. However, for women, there are a lot of different treatment options that can vary between surgery to several therapies and a combination of them.

The treatments can be classified as local treatments or systemic treatments. The local treatments are applied on early cancer stages, treating tumors without affecting other parts of the body. When the cancer has already spread to surrounding tissues, systemic treatments are preferred in order to treat tumor cells on different areas of the patient's body. Some of the examples of local treatment are surgery and radiotherapy, while systemic treatments are composed by chemotherapy, hormone therapy or targeted therapy. This types of treatments may be combined to reach better results leading to two different types of therapy: neoadjuvant therapies and adjuvant therapies, that combine local and systemic treatments in order to treat a patient. The neoadjuvant therapy consists on applying systemic treatments before the surgery to reduce the carcinoma size. This allows to remove a smaller portion of the tumor during surgery and can be applied on cases that were considered inoperable due to the increased tumor size. The adjuvant therapy consists on

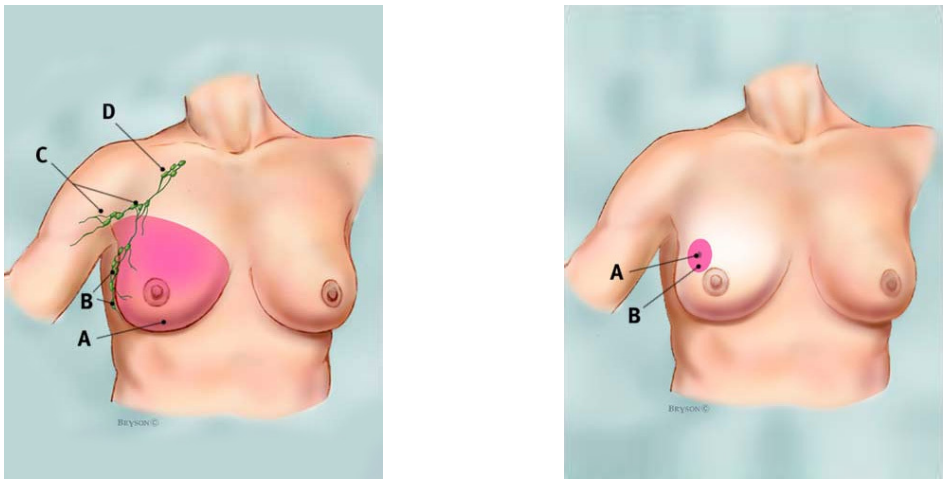
[7] https://radiopaedia.org/articles/breast-mri
[8] https://radiopaedia.org/articles/breast-density

complementary treatments applied after the surgery in order to eliminate cancer cells that may have spread previously to the surgery or eliminate cells not removed during the surgery [DD16].

As mentioned before, surgery is the most frequent option to treat breast cancer. However, it may be done for other reasons either than removing the tumor and the surrounding tissue, listed below:

- Perform biopsies on sentinel lymph nodes, in order to find if the cancer cells had spread to the axillary lymph nodes;
- Breast reconstruction, to restore the breast shape after removing the tumor;
- Relieve symptoms of advanced breast cancer.

The most common types of surgery are Mastectomy and Breast Conserving Surgery, represented in Figure 2.4. Both are performed in order to remove the tumor and the surrounding tissue on the patient's breast.



(a) Mastectomy (b) Breast Conserving Surgery

Figure 2.4: Breast Cancer Surgery [9]

- **Mastectomy**

  Even though the different approaches within the Mastectomy: radical mastectomy, modified radial mastectomy, simple mastectomy and skin-sparing mastectomy, all the above consists on the extraction of the entire breast tissues and some of the surrounding tissues. In the case of radical mastectomy, the pectoral muscle and all the axillary lymph nodes are also removed. Despite of the remotion of a great volume, the mastectomy do not remove tissues beyond the clavicle, the inframammary fold (above the rectus sheath), the midline of the sternum and the anterior border of the latissimus dorsi [Wit04]. Less radical surgeries have

[9] http://www.breastcancer.org/treatment/surgery

been proved as effective as this. Nowadays, mastectomy is only performed for patients with large tumors that have spread into the pectoral muscle.

- **Breast Conserving Surgery**

  Breast Conserving Surgery (BCS), as known as partial mastectomy or lumpectomy, is a less extreme way to surgically remove the tumor. This procedure aims to remove the least possible amount of breast tissue, containing only the tumor and a small portion of the healthy tissue surrounding it. BCS is much less invasive than Mastectomy and with the same effectiveness and survival rates when radiotherapy is used as complementary treatment. Even though some women prefer the mastectomy, fearing the recurrence of the cancer, studies have proven that mastectomy do not provide any better result regarding the cancer treatment and recurrence than BCS, and the number of patients choosing this option as treatment has increased.

2.4 Impact of Breast Cancer Treatment

2.4.1 Influence on the Aesthetic Outcome

Several studies have been made in order to understand what parameters and how they influence the aesthetic outcome of BCS. These studies have evaluated parameters such as the age of the patient, her body mass index, the existence of palpable tumors, the location, volume and weight of the tumor, the axillary and breast incisions during surgery, among other factors.

It is also known that radiotherapy and chemotherapy have influence on the aesthetic result; however, all the studies regarding the adjuvant therapies influence are merely based on empirical experiences on patients.

According to [FGH+14] previous studies have concluded that, generally, when the tumor is located on the inner quadrants of the breast, the treatment results in a poor aesthetic outcome. The location of the tumor on the nipple areolar complex leads to the most unfavorable aesthetic outcome. Unlike the location, the tumor weight has not been proved as predictive of a bad aesthetic outcome, in spite of tumors with less than 50g tend to lead to good aesthetic outcomes according with the aesthetic evaluation metrics presented in [JVM+12].

Concerning, the breast incision and besides of the inconsistent findings, some authors associate better aesthetic outcomes with radial and circular incisions, and worst outcomes with the periareolar one [FGH+14]. Other studies point aspects like the result of mechanical forces such as gravity, breast tissue constitutive law distribution, inflammation due radiotherapy, the internal stress generated by the healing process and angiogenesis as factors on the breast shape after Breast Conserving techniques [VEH+16]. The aspect of the breast may change significantly during the healing process that can take as long as two years due to alteration regarding the tissue composition, stiffness and volume.

2.4.2 Quality of Life

Changes of the breast's size or shape or even the loss of it have a significant impact on the psychological and social life of a patient. The loss of breast as a symbol of femininity has may lead to a decrease of the patient's self-esteem, negative body image, social isolation and communication and relationship problems. The psychosocial stress and the physical burden may result in a reduction of the patient's opportunities in life and increase their social rejection. Common symptoms of the breast cancer treatment visible on a wide number of patients are anxiety, depression, fatigue, pain, difficult in concentration, sexuality concerns and self-blame [AAAAAR+14]. Despite of the sparsity of studies regarding the influence of breast treatment surgeries on women's body image and sexuality, the most recent ones consider BCS result on a best preservation of the woman's body image and more comfort about their sexuality [Row00].

2.5 Summary

Given the importance of the visualization of a treatment outcome on a breast cancer treatment and the impact it has on the QoL of the patient, the ability of predicting the aesthetic outcome of a surgery is a valuable aspect.

Currently, this is shown with 2D drawings and pictures of similar cases, and in some cases with generic 3D models. A patient-specific prediction will allow to minimize the doctor-patient misunderstandings, to find the optimal treatment, to decrease the fear the patients stood against surgery and to select the most desired outcome of the surgery. In order to achieve this, it is important to be able to manipulate a 3D breast model, and define the transformations that each treatment may led on the patient's breast. As outcome, the patient may be able to see own breast deformed in a real time system.

The next chapter describes the models that are being used so far and how a breast model may be represented as well as some existing simulations of the breast surgery.

Chapter 3

3D Modeling of the breast

In order to achieve a prediction of the aesthetic outcome after a BCS the best option is to describe the breast using 3D models. Some models can be reached with several techniques, but all of them retrieve a model with some compression or mechanical force present.

Due to the importance on medical imaging to find a smooth surface that fits a set of 3D unstructured data to describe and represent anatomical structures, many alternatives have been found over the years. Being a complicated problem and due to the variability and difference between the shape of human breasts, a lack of work on this field was led to [DD16].

The present chapter starts by presenting several data acquisition techniques, in section 3.1, that could provide the information for generating 3D models. The data may be represented using for example parametric and deformable models presented in section 3.2 and section 3.3, respectively. At the end of the chapter some existing frameworks used for breast augmentation and plastic surgery are referred in section 3.4, such as the methods that they use in order to acquire and model 3D data.

3.1 Data Acquisition

Despite of all the several ways to describe and represent the shape of breast, the 3D imaging yields more information than multiple conventional photographs. Given this, 3D models are the best and more realistic way to evaluate the shape and size of the breast, its symmetry, contour, volume and surface area [KSR+08].

Multiple attempts have been done to represent breast as a 3D object. Those attempts have started by using Magnetic Resonance Imaging (MRI), Computed Tomography (CT) and 3D surface imaging systems. Bücking et al. in [BHR+17] describes the use of CT and MRI data to generate 3D models of anatomical human parts in order to 3D print the obtained models. In this case a pre-processing of the data was done considering two main steps: image segmentation, where the image was labelled and partitioned into several areas and regions ignoring the noisy regions

of the image; and a mesh refinement, repairing and smoothing the models' discontinuities. The authors also mentioned that CT was used instead, when segmenting structures with low or high densities. MRI are best used in soft tissues due the high contrast on this cases.

An alternative to retrieve 3D models of the patient's torso was proposed in [CMZO14]. This approach used a low-cost depth sensor (Microsoft Kinect) to acquire several views of patient's torso in order to perform a point cloud registration of the breast. The point cloud registration process is subdivided into two parts: coarse registration and fine registration. In order to generate the point cloud that would serve as input in the coarse registration, the raw RGB-D data acquired by the sensor were pre-processed. The pre-processing started by segmenting and then filtering the depth image in order to remove the noise on the edge and silhouette of the object. Given the retrieve point cloud, a Tessellation-based coarse registration uses depth data to align the point clouds. The alignment was done by a pose estimation, to reduce the initial misalignments; a keypoint selection to identify some correspondences between different point clouds; and a correspondence estimation and validation to find the better coarse alignment.

According to [AMS76], the acquisition of MRI data consist of several MRI axial slices of the breast and ensures the 3D visualization of the patient's breast. In the cases where the acquisition of MR images is not axial, it can be converted posteriorly. Thereafter and despite of the semi-automatic segmention of the MRI data contours, the images required a manual segmentation in order to differ parenchyma, fat and lesion tissues. Based on the segmentation results and the defined contours it is possible to calculate a few reference points and then generate a 3D computational mesh. The generated mesh is represented by parallel planes limiting the breast's contour.

3.2 Parametric Models

Once the raw 3D data are obtained and due to the need to easily manipulate it, depending on its application, the information acquired may need to be represented or transformed in any other type of 3D representation. In the case of medical application such as representing human organs, the parametric models are widely used. An application of parametric models was described in [VU98] when representing the left ventricle of the heart. These models have as advantage using superquadrics parameters allowing to represent objects with rounded edges or corners that may resemble a wide variety of human organs.

The problem of fitting supperquadrics to 3D unstructured data, what solved by the use of some robust and fast methods as described below. A superquadric refers to different sets of superellipsoids, supertoroids or even one or two pieces of superhyperboloids. And despite of the general use of superellipsoids, presented in Figure 3.1 the other objects can also be used in order to describe different shapes.

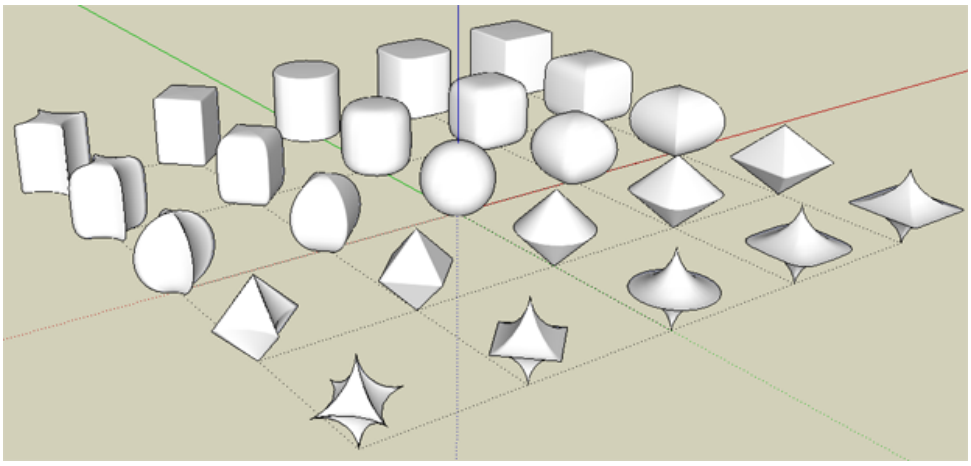


Figure 3.1: Examples of superellipsoids [1]

A superquadric is obtained by the spherical product of two 2D curves. In the case of a superellipsoid, it is described by the following equations.

$$\left(\left(\left(\frac{x}{a_1}\right)^{\frac{2}{\varepsilon_2}}+\left(\frac{y}{a_2}\right)^{\frac{2}{\varepsilon_2}}\right)^{\frac{\varepsilon_2}{\varepsilon_1}}+\left(\frac{z}{a_3}\right)^{\frac{2}{\varepsilon_1}}\right)^{\frac{\varepsilon_1}{2}}=1 \tag{3.1}$$

The parameters a_1, a_2 and a_3 are scaling factors corresponding to the coordinate axis, and ε_1 and ε_2 represent the squareness of the original superellipsoids.

A smaller ε represents a superellipsoid similar to a square, ε equal to 1 represents a circle and ε equal to 2, a shape with a flat bevel and a superellipsoid with larger ε defines a pinched shape.

In order to pose the superquadric on an axis system, 6 more parameters are required.

$$Xw = T.Xs, T = \begin{bmatrix} R & t \\ 0 & 1 \end{bmatrix} \tag{3.2}$$

R is a 3x3 matrix and t a 3x1 matrix representing the rotation and translation, in relation to the referrals' origin, of the superquadric respectively.

As soon as the most adequate superquadric is chosen, it must fit the 3D data that we want to represent. Despite of the good global approximation of a shape, superquadrics are too limited when representing more complicated surfaces. This problem usually occurs due to the symmetry of the superquadrics. A possibility to overcome is the utilize deformable models presented on section 3.3 or its application of the parametric models refereed in [PCO14].

[1] http://regular-polygon.com/plugins/superellipsoid/

3.3 Deformable Models

Representing the 3D data of the patient's breast using Deformable models will allow generating and manipulating complicated curves and surfaces. The deformations that would be applied in order to generate that kind of complex data can be categorized in two different methods Physical Methods and Non-Physical Methods. While the Non-Physical Methods manually manipulate and deform the objects by adjusting one or more parameters of the shape (that describe the more simple object), the Physical Method relies on the modification of the physical properties of the object through the application of external forces. In the case of Physical methods, the material's properties of the object also impact the deformation of the object [GM97, DD16].

3.3.1 Non-Physical Models

As mentioned before, the non-physical methods to deform object are done recurring to the alteration of model parameters. Widely used ways to represent curves defined by vectors of control point vary between Bezier curves, B-splines or non-uniform rational B-splines (NURBS).

The abovementioned approach in [PCO14] and [VU98] is based on the application of Free Form Deformation (FFD). The FFD allows to define the deformations by adjusting the space where the object lies and not by changing its control points. Another advantage of this approach is that the same deformation can be applied to the different models simultaneously.

On this specific approach the model is considered to be embedded in a box that can be changed in order to twist, bend or taper the model on its interior. Figure 3.2 shows the object embedded on the box of control points in 3.2a and the result of the object deformation in 3.2b. To accomplish this, the FFD formulation shall be done in the two following steps:

1. Compute the local coordinates of the object points in the frame defined by the box of control points.

$$X = X_0 + sS + tT + uU, \tag{3.3}$$

where s, t and u are given by:

$$s = \frac{S.(X - X_0)}{S.S}, t = \frac{T.(X - X_0)}{T.T}, u = \frac{U.(X - X_0)}{U.U}, \tag{3.4}$$

X represent each point of the objects by the coordinates (s,t,u) and the box where the object is embedded is represented by the vertex X_0 and the box edges (S,T,U).

The point X is inside of the box if and only if s, t and u have all values between 0 and 1.

The size of the embedding box is given by the parameters a_1, a_2 and a_3 of the superquadric and its rotation is given by the coefficients of the rigid transform φ, θ, ψ, t_x, t_y and t_z.

The volumetric grid of the box's control points $(l+1)\ (m+1)\ (n+1)$ can be described by:

$$\begin{cases} x(P_{ijk}) = a_1\left(1-2\frac{i}{l}\right) \\ y(P_{ijk}) = a_2\left(-1+2\frac{j}{m}\right) \\ z(P_{ijk}) = a_3\left(1-2\frac{k}{n}\right) \end{cases} \tag{3.5}$$

At last, the space alterations that the model will be put through may be represented as:

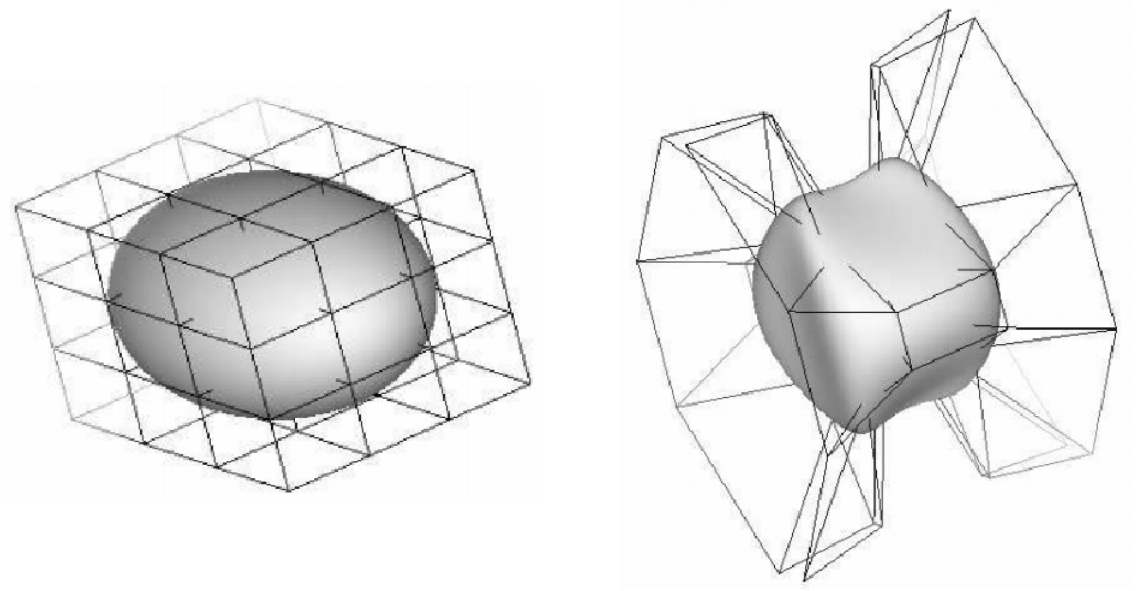
$$X = BP, \tag{3.6}$$

where X is a matrix NP x 3 (NP: number of control points = (l+1)(m+1)(n+1)) with the coordinates of the model points, B is the deformation matrix ND x NP (ND: number of points on the object) and P the NP x 3 matrix which contains coordinates of the control points P_{ijk}.

2. To achieve the best fitting of the model to the data we intent to represent, the displacement field must be reduced. This displacement refers to the distance between the model and the data points we will represent.

   We are changing position of control points to fit X to the target model. Note that we are not fitting control points to sth.

   As soon as the best fitting of the model is find, by changing the position of control points in order to make X fit the target model, the position of point X of the object may be computed.



(a) Box of control points with embedded object

(b) Object of the deformed box

Figure 3.2: FFD deformation [VU98]

3.3.2 Physical Models

A well-studied physical model is mass-spring method which is used in modeling facial expressions. The proposed methodology described in [GM97] uses three distinct layers of tissue in order to deliver a mesh of mass points corresponding to the dermis, a layer of fatty tissue and a layer of muscular tissue. The same approach has been adopted in the context of breast surgery [PIS+13], where a volumetric tetrahedral mesh representing the breast was computed from a semi-automatic segmentation procedure. Then the mesh was deformed based on the mass-spring model: the spring's rest length and stiffness were estimated and then applied to the uncompressed breast model in order to deform it to the real compressed one. Although being easy to construct and allow to deform the objects in more ways than other physical methods, mass-spring finds difficult to model incompressible volumetric objects or unbendable surfaces.

Another method with a great variety of applications is Finite Element Models (FEM). In contrast with the mass-spring method, FEM is more accurate, requiring a much larger computational power and being a very time consumption process. In a FEM, the object is divided into several elements joined by discrete node points. The desired deformation function is then applied to each element in order to find an approximation that satisfies an equilibrium expression relative to the intended deformation. The type of elements that are used to form the model are chosen according to the properties of the object, the trade-off between the computational power and the required accuracy. One of many examples of the application of FEM was described in [KM04] where simullating the interacting between the soft tissue of a human hand and a deformable object. Considering the apllication of FEM in breast surgery, it has been used in [VEH+16], whose proposed methodology will be further detailed.

Vavourakis et al. [VEH+16] proposed a surgical simulator based on a FEM. In order to simulate the wound healing effect described in [VEH+16] data was gathered though MRI as described in [AMS76]. The used mesh is constituted by two distinct types of isoparametric elements, shown on Figure 3.3:

- Solid 4-node trilinear isoparametric elements - used to represent the breast tissues except for the skin;

- 3-node triangular isoparametric elements - used to represent the surface of the model: the breast's skin.

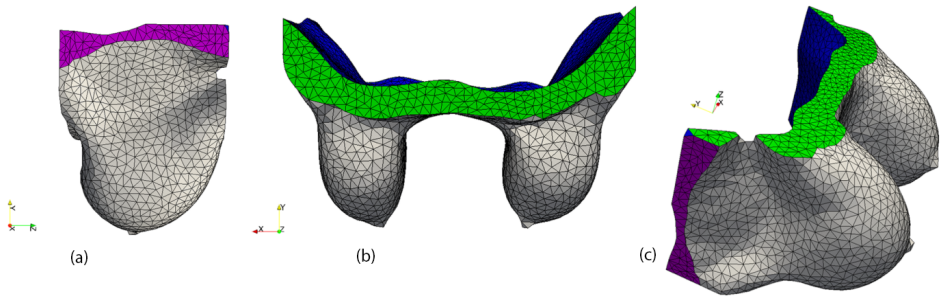


Figure 3.3: **3D mesh with isoparametric elements** (A) Lateral View (B) Caudal View (C) 3D View [VEH+16]

In the obtained 3D mesh, a material is assigned to the elements that represents it on the mesh. This assignment is based on the different types of tissue: fat, parenchyma and damaged. Each type of tissue on the mesh will be assigned it a different type of materials.

Vavourakis et al. [VEH+16] described the implementation of a surgical simulator based on Multiscale FEM, where two concurrent simulations were performed: a wound healing simulation and a biomechanical simulation. This implementation is represented in Figure 3.4, whose steps are described below:

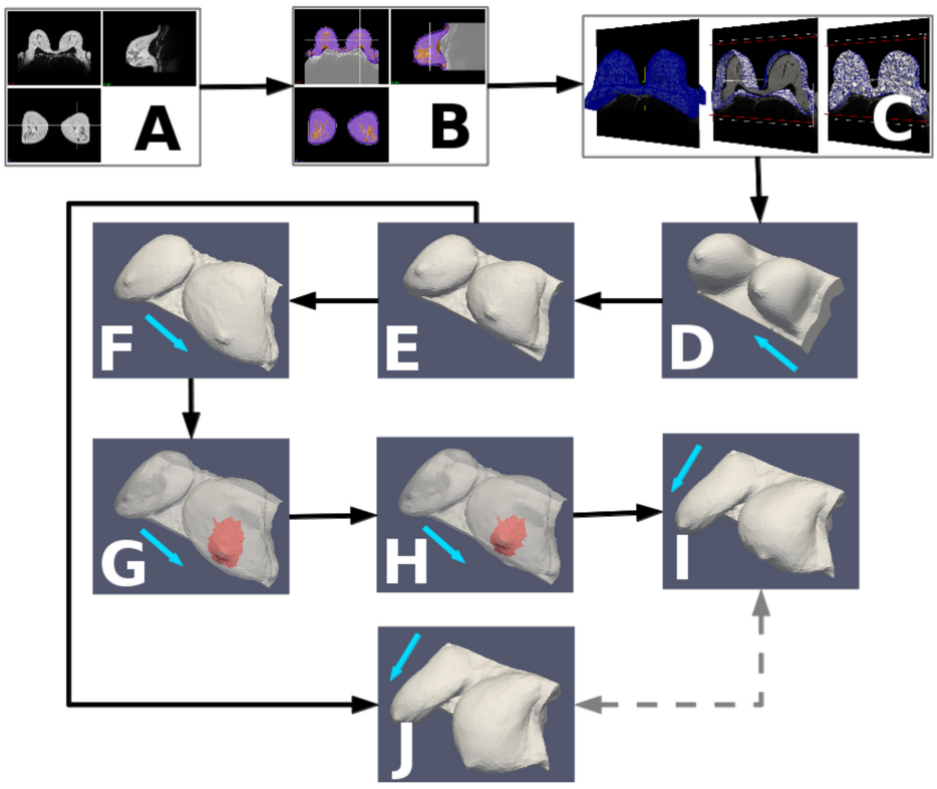


Figure 3.4: Example of computational process for surgical simulator [VEH+16]

- In step A, the MRI data and the computed-tomography (CT) data are acquired;
- In step B, the data is segmented, differing the adipose and fibroglandular tissues of the breast;
- In step C, the generation of the surface and volumetric meshes of the patient's breast is possible through the data acquisition and segmentation referred in the previous steps;
- In step D, the models are prepared for the input of the Finite Element solver;
- The retrieved data from MRI are represented in prone since they are stressed by the gravity force. In order to apply mechanical finite element models, the step E is required to remove the gravity effect on the model, computing the a gravity unloaded model;
- In step F, the gravity unloaded model of the breast is converted into a supine geometry;
- In step J, the gravity unloaded model of the breast is also used to predict the breast's geometry in upright;
- In step G, the tumor position is identified through determining the incision lines and the outline of the excised tissue. Consequently, the elements inside the outline of excised tissue are labeled as damaged tissue;
- Step H, performs the wound healing simulation resulting in the wound contraction and the estimation of the post-surgery breast shape;
- Finally in step I, the effect of gravity is re-applied on the breast shape, retrieving the model in a stand up position, or upright geometry.

The proposed surgical simulator that estimates the wound healing and the pos-surgical shape of the breast relies on the two mathematical models described below:

- Wound Healing and Angiogenesis Model

  This mathematical models is based on the cell density, the concentration of biochemical agents responsible for the mitosis regulation, the density of the microvascular density, the nutrient and oxygen levels and the agent that regulates vascular spouting in order to compute the changes on the breast shape during the healing process. It also takes into consideration the increase of the inflammatory response and the stimulation of the immune system.

- Soft Tissue Biomechanics Model

  Also responsible for the configuration of the breast geometry, the Soft Tissue Biomechanics model takes into account the breast tissue's mass density and the body force vector. Within this model, the stress distribution is calculated considering the passive stress of the tissues'

mechanical deformations and the active stress from the tissue recovering during the wound healing process.

3.4 Existing Frameworks

The greater portion on this field focuses on breast augmentation and plastic surgery and have originated some software able to simulate deformations on the breast tissue. Being the most known:

- Crisalix [2]

  Crisalix is a web based application based on 2D photographs with a range of implant types, sizes and surgery techniques available. Figure 3.5 shows Crisalix software, that allows to simulate breast enlargement or reduction, breast lift, the application of silicone implants, implants revision, scars, breast reconstruction and fat transfer.

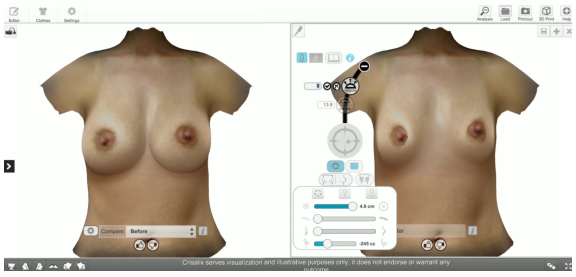


Figure 3.5: Crisalix interface [2]

- Sculpt My Dream [3]

  Sculpt My Dream is a property platform of Vectra3D, and uses six distinct cameras to reconstruct a virtual model of the patient's torso. The simulation relays on a variety of implant sizes and a list of manufacturers while is able to correct some asymmetry of the breast. Even though the good estimation and the similarity between the software simulation and the procedure, it can only be applied to plastic surgery. Figure 3.6 shows Sculpt My Dream interface.

[2] https://www.crisalix.com/en
[3] http://www.sculptmydream.com/

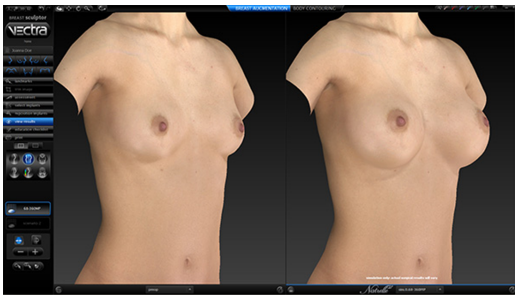


Figure 3.6: Sculpt My Dream interface [3]

- Axis Three [4]

  Axis Three uses a property scanner to capture 3D images of the patient's torso. It simulates alterations on the face or breast of the patient. In the specific case of breast simulation, it is based on the manufacturers implant, the location of the implant (beneath or above the pectoral muscle) and the tissue's elasticity. Figure 3.7 shows Axis Three interface.

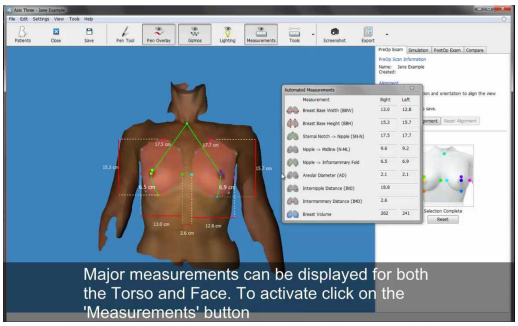


Figure 3.7: Axis Three interface [4]

In spite of the reliability of these 3D scanner, the utilization of several cameras are demandingly expensive and while it requires complicated procedure to perform 3D reconstruction of the patient's torso.

3.5 Summary

The patient's 3D models, including torso, or individual organs, have been commonly used in applications of both medical tools. In this chapter, a brief descriptions of recent methodologies have been explained as well as their medial applicastions.

Concerning the representation of humans breast, most of the studies focused on the application of the models for plastic surgery. More recently, FFD was used in order to represent the breast's shape; and FEM researched in order to simulate wound healing and transformations between different geometry configurations of the breast.

[4] http://www.axisthree.com/welcome

Despite of the existence of a surgical simulator, time consumption and computational power requirements of FEM makes it unviable to be used in real-time systems. Machine learning techniques will be applied in order to predict the same deformation caused by the BCS in order to overcome the problems of the FEM approach. This methodology is detailed in chapter 5 and its results are presented in chapter 6.

Chapter 4

BCS planning tool

A BCS tool, having as main objective to assist health professionals to understand the surgical outcome, requires the definition of the tumor on the specific breast of the patient. The same tool will also be used to create the dataset as explained in section 5.1. The tumor location and size will be used to predict the aesthetic outcome of a BCS considering the shape of the patient's breast and the specific parameters of the tumor, using the obtained prediction models described in chapter 5.

In this chapter the Functional such as Actors, Functionalities, Use cases and User stories and non Functional requirements of the software are presented, as well as some considerations concerning the interface's design and the application flow.

4.1 Functional Requirements

The Functional requirements describes the system and its functionalities and components such as the actors and the way they interact with the system.

4.1.1 Actors

4.1.1.1 Health Professional

Both surgeons and radiologists are considered as health professional and are allowed to perform any of the functionalities presented on the tool.

4.1.2 Functionalities

The developed tool consists of the following functionalities:

- Loading Breast Model
- Exporting Tumor and Excision Information, considering the breast's density

- Visualizing model:
  - Zoom model
  - Change model Point Size
  - Orthogonal views
- Locating nipple
- Breast Division:
  - Show / hide breast's quadrants
- Defining tumor :
  - Tumor position, by either selecting a point or randomly choosing a position within the defined region
  - Tumor size
- Defining excision:
  - Excision radius
- Visualizing information:
  - Tumor position - quadrant
  - Breast's Laterality
  - Breast's Volume
  - Tumor radius
  - Tumor volume
  - Excision Volume
  - Excision / Breast volume ratio
- Undo
- Redo
- Reset

These functionalities will be further detailed on the use cases diagram and the user stories table.

4.1.3 User Stories

The system's user stories are described in Table 4.1.

Table 4.1: User Stories of BCCT planning tool

ID	Name	Priority	Description
US001	Load Model	High	As a User, I want to load a specific breast's model of a patient.
US002	Model Visualization	High	As a User, I want to rotate, zoom or pan the loaded model for visualization.
US003	Change orthogonal View	High	As a User, I want to visualize the breast model on an orthogonal view (front, top, bottom, back, lateral).
US004	Define Nipple Position	High	As a User, I want to be able to define the nipple position on the point cloud.
US005	Breast division	Medium	As a User, I want to show or hide the breast quadrants (Upper Outer, Upper Inner, Lower Outer, Lower Inner).
US006	Zoom in / Zoom out	Medium	As a User, I want to Zoom in or Zoom out the breast's model.
US007	Change point size	Medium	As a User, I want to Increase or Decrease the point cloud's point size.
US008	Define the tumor position	High	As a User, I want to define the tumor's position, either by picking a point or randomly choose a point within the defined region.
US009	Define the tumor size	High	As a User, I want to define the tumor's size.
US010	Draw the tumor sphere	High	As a User, I want to see the tumor sphere drawn over the breast's point cloud.
US011	Define the Excision radius	High	As a User, I want to be able to define the margin between the excision cylinder and the tumor's radius.
US012	Draw the Excision's cylinder	High	As a User, I want to see the excision's cylinder drawn over the breast point cloud.
US013	View the tumor quadrant	High	As a User, I want to visualize in which breast's quadrant the tumor is located in.
US014	View the breast volume	High	As a User, I want to visualize the selected breast's volume.
US015	View Tumor radius	High	As a User, I want to visualize the calculated tumor radius.
US016	View the tumor volume	High	As a User, I want to visualize the tumor's volume.
US017	View the Excised volume	High	As a User, I want to visualize the volume to be excised.

Table 4.1: User Stories of BCCT planning tool

ID	Name	Priority	Description
US018	View Excision/Breast volume ratio	Medium	As a User, I want to see the percentage of the breast volume to be excised.
US019	Export excision and tumor point cloud information	High	As a User, I want to record the model points that were considered the tumor or the excision portion.
US020	Reset	Medium	As a User, I want to restore the state of the system, reloading the breasts' models to their initial state.
US021	Undo	Medium	As a User, I want to restore the system to the previously state, ignoring the last step.
US022	Redo	Medium	As a User, I want to perform the previously undone step.
US023	Hide the tumor sphere	Medium	As a User, I want to hide the tumor sphere previously drawn.
US024	Hide the excision cylinder	Medium	As a User, I want to hide the excised cylinder previously drawn.

4.1.4 Use cases

Figure 4.1 describes the possible interactions that a user can have in the tool. Note that simple interactions were ignored.

4.1.5 Functional Constraints

The developed tool have some constraints, due to the representation of the breast model as a point cloud. Those constraints are presented below:

- When selecting both the nipple and tumor position, must be a point from the model's point cloud;
- The calculation of the breast, tumor and excision volumes are calculated through approximations;
- When defining the tumor's sphere and the excision's cylinder, the breast boundaries are not taken into account. This way, the defined polygon will be drawn disregarding the point clouds' limits.

4.2 Non Functional Requirements

These are the requirements that are not crucial for the normal function of the application but enhance the user's experience.

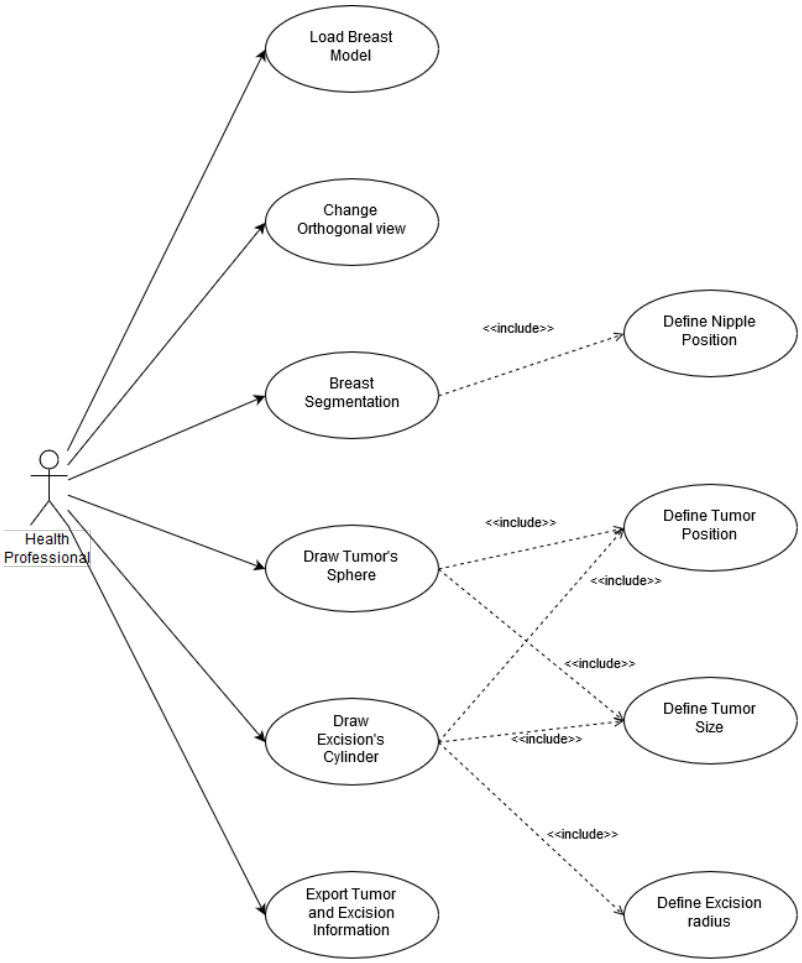


Figure 4.1: Use cases for the BCS planning tool

- **Interface:** The user interface must be intuitive allowing the user to access all the functionalities with the minimal interaction with the tool in a logical way. To enhance this process many of the functionalities despite of being always visible are only enable when the system has enough information to perform it. In some cases the system warns the user to add necessary information before performing the task;

- **Maintenance:** The tool was developed in a way to easily modify the implemented functionalities of the system;

- **Expandability:** The tool was developed in a way to easily extend or add additional functionalities to the system;

- **Efficiency:** The tool performs all the system's function on a short time.

4.3 Application Flow

The flow presented in Figure 4.2 shows the required steps to perform the BCS planning. All the tasks not presented on the flow are done at any moment after loading the breast's model.

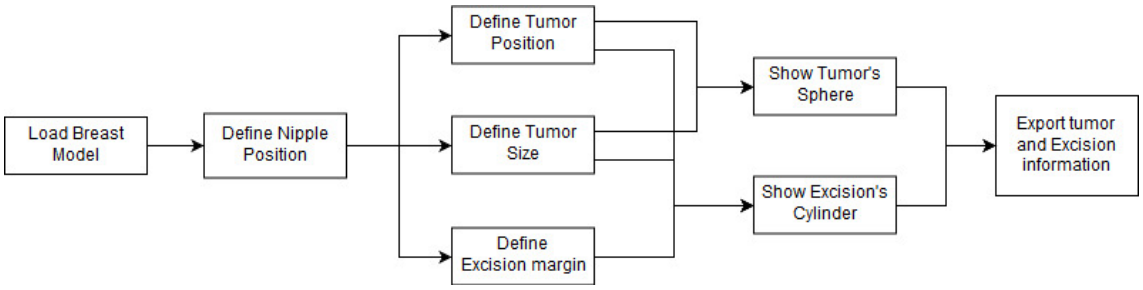


Figure 4.2: Tool flow with the required steps to perform the BCS planning

4.4 Development

The frameworks used for the development of the described application are explained below and were implemented using C++. A few frameworks such as Qt and VTK were also used in order to simplify the tool's interface development.

4.4.1 System Documentation

All the developed codes were written clearly and in a straightforward way in order to allow a intuitive and fast reading of each function and described object.

4.4.2 C++

All the required functions where declared on the header files and respectively implemented on the correspondent *cpp* file. Besides the utilization of this types of file, there is a main file responsable for the execution of the interface and all the functions associated with it, and a *ui* (user interface) file where the interface was designed.

4.4.2.1 Code comments

The code comments, as a way of code documentation, are only used when necessary decreasing the amount of visual barriers, often associated with mathematical formulas or required calculations on the 3D space.

4.4.3 Used Frameworks

4.4.3.1 Qt

Qt allows the development of multi-platform applications and interfaces based on C++ in a simple and fast way [1].

4.4.3.2 VTK

The Visualization Toolkit (VTK) is an open-source software system used for 3D computer graphics, image processing, and visualization that consists of a C++ class library including several interpreted interface layers [2]. VTK integrates other frameworks such as Qt, also used for the development of this application.

4.4.3.3 Boost library

Boost library is a C++ set of libraries that allows an easy utilization of linear algebra, image processing and multithreading. These libraries are required for the utilization of VTK and Qt frameworks.

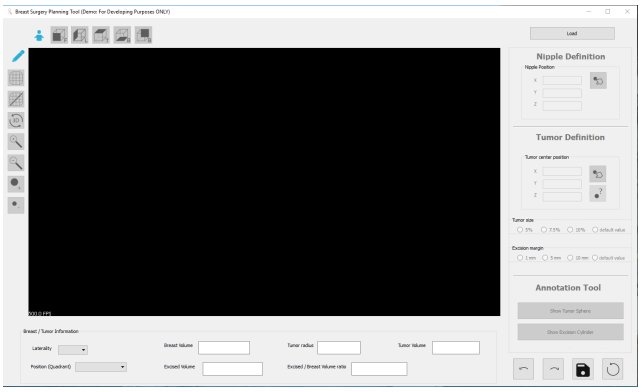
4.5 Interface

Figure 4.3 shows the interface of the main functionalities of the tool. When launching the tool, the initial window will be displayed, where it is only possible to load a breast's point cloud (a). After loading the breast's model, the nipple's position feature is enabled. The breast's point cloud is also displayed on the visualization area. A selection point view for the nipple position is presented in (b). For the tumor position a similar view will be displayed with the breast seen from a frontal position. After defining the nipple position, the breast can be divided into quadrants. This can be seen through planes drawn over the breast in (c). In order to do the tumor definition, tumor position is selected as well as a tumor size. If one of the fields is not checked, a warning window will be displayed (d). When all the fields are completed, a tumor's sphere can be drawn and displayed over the breast on the visualization area (e). Last, but not least, when selecting an excision margin, the excision's cylinder is drawn (f).

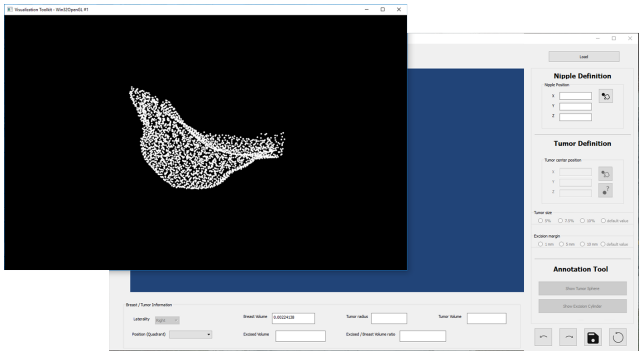
Regarding the interface, this was created considering the target group and adapted according to its propose. One of the aspects that were took into consideration was the used icons. Due to the lack of materials guidelines and icons to represent some of the intended actions, some icons used in the tool were created in order to overcome that lack and the ones that already existed were adjusted to provide an overall consistency among all the tool. Other human-computer iteration principles were considered as the representation of already clicked buttons, and the pop-up of warning and dialog messages to inform the user about any error or task completion. The tool's

[1] http://doc.qt.io/
[2] http://www.vtk.org/documentation/

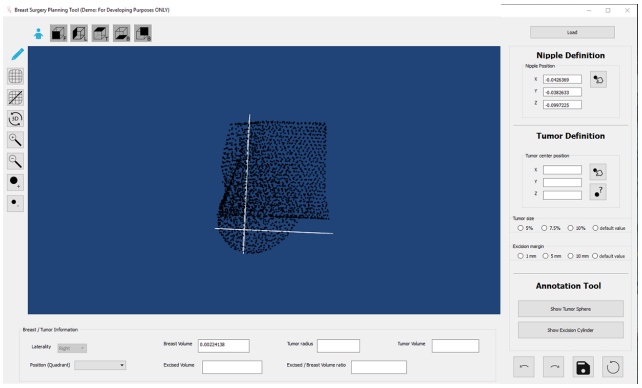
interface is displayed in Figure 4.4, where the interaction panels are delimited and coupled with a label.



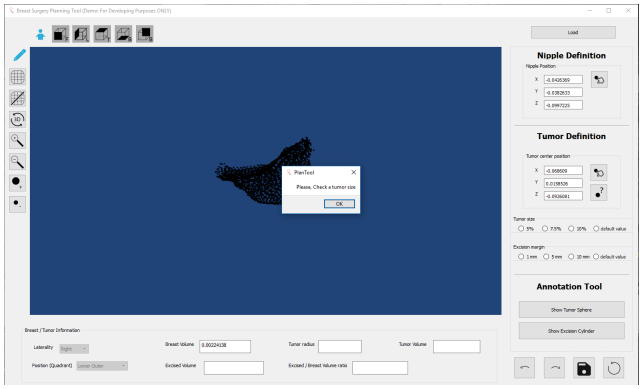
(a) Initial view



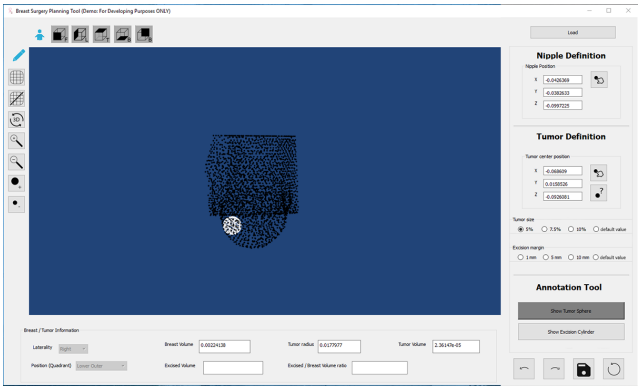
(b) Point Selection View



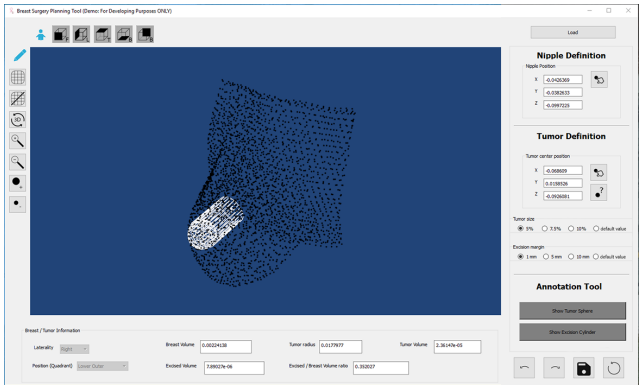
(c) Breast Segmentation View



(d) Warning Alert



(e) Tumor's view



(f) Excision's view

Figure 4.3: Interface of several views of the BCS planning tool

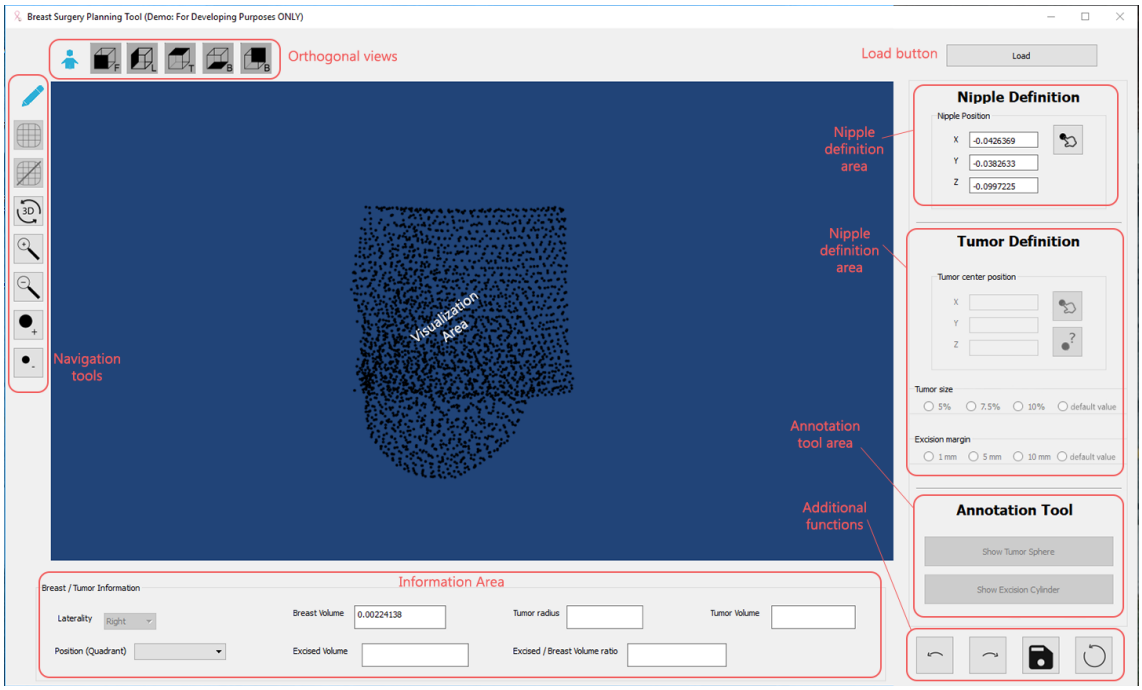


Figure 4.4: BCS Tool Interface

4.6 Summary

In this chapter the developed BCS planing tool was presented as well as both functional and non functional requirements, application flow, the frameworks that were used. Also some considerations regarding its development and implementation are described as well as some interfaces of the application.

One of the most valuable functionalities that the tool can be equipped with is to allow the simulation and further visualization of the breast's deformations predicted by the models described in chapter 5.

Chapter 5

Methodology

In this chapter all the process that led to achieving the main goal of the thesis, predict the deformations caused by BCS is described in detail. In order to plan such process, the study on breast cancer and how to represent the Breast as a 3D model will be used, as well as the developed tool described in chapter 4.

The applied methodology is divided in 3 fundamental parts: a dataset preparation to feed the following steps, the application of Machine Learning in order to predict the deformations on the patient's breasts caused by the BCCT, and the validation of the obtained results through several metrics. Those three steps are represented on Figure 5.1 and are detail through the several sections on this chapter.

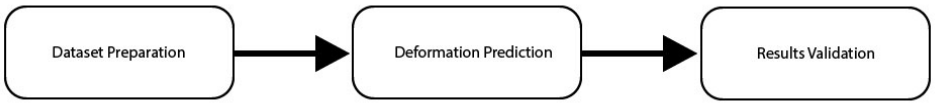


Figure 5.1: Flowchart of the applied methodology

5.1 Dataset Preparation

In order to further apply machine learning techniques to predict the deformations on the patient's breast, a semi-synthetic dataset was prepared. Despite of existing a few datasets with 3D models of the breast, they are constructed based on a generic shape of the breast with some deformations applied to it. Although the great amount of studied deformations, they are still unable to represent the variability and diversity that may be found in a semi-synthetic dataset. This created semi-synthetic dataset contains 3D breast models representing the patient's breast before and after the surgery. The pre-surgical models in the dataset are based on the real data obtained though MRI data

of a few patients. The pos-surgical models are generated by taking as parameters the hypothetical tumor's location and volume and the breast's density. Considering different parametrization, each pre-surgery model of the dataset defer from other regarding the following features:

- breast's density,
- breast's volume,
- breast's laterality,
- tumor's position
- tumor's volume.

The pos-surgery 3D models of the breast are obtained though a biomechanical simulation of the wound healing based on the pre-surgical models generated through the patients' MRI data. The BCS planning tool described in chapter 4 was used to generate the hyphothetical tumor's location and volume. The pre-surgery models will be used to simulate the pos-surgery models alongside with the hyphotetical tumor's location and size through a biomechanical healing simulator described in [VEH+16]. Figure 5.2 illustrates the workflow of the steps pursued for preparing the dataset.

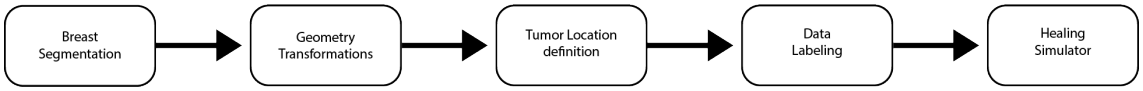


Figure 5.2: Flowchart for the dataset generation

Initially, it is necessary to segment the breast outline from background of the MRI data for each initial real patient. Further, a 3D model is reconstructed from the segmented data. However, since the MRI is taken as the patient is lied in prone position, it is necessary to apply a geometry deformation to represent the breast in a supine configuration. With the proper geometry, the tumor must be defined in the specific breast, that therefore will be labelled according with the healthy and damaged portions of the breast. At last, the wound healing simulation will be performed in order to generate the pos-surgical model of the breast.

5.1.1 Breast segmentation

As described before, the pre-surgical 3D models of the patients' breast were constructed based on MRI data segmentation. The breast segmentation is performed by annotating both the breast and the pectoral muscle. The annotation included both left and right breast with the respective nipple and had as lateral boundaries the left and right latissimus dorsi. The pectoral muscle included both the right and left major and minor pectoral muscles. Figure 5.3 depicts the anatomical structure mentioned during segmentation. The result of a patient annotation used to generate the breast's point cloud is shown in Figure 5.4.

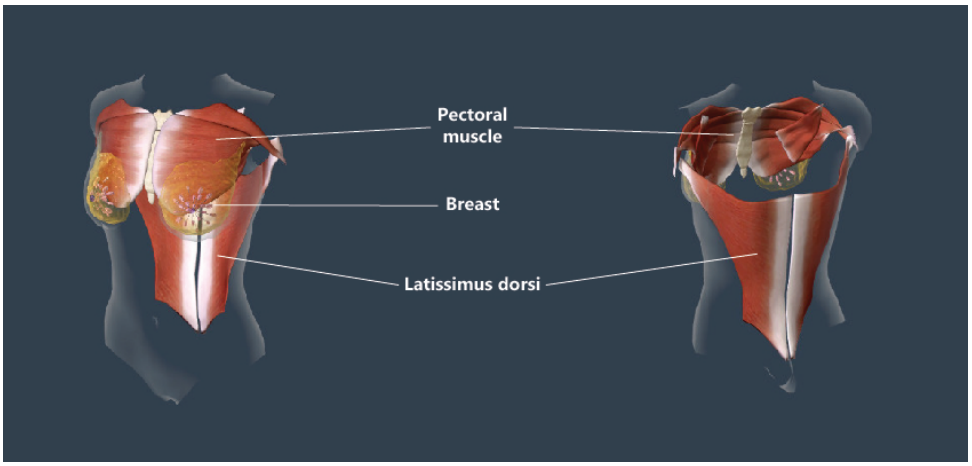


Figure 5.3: Representation of the Breast, Pectoral muscle and Latissimus Dorsi

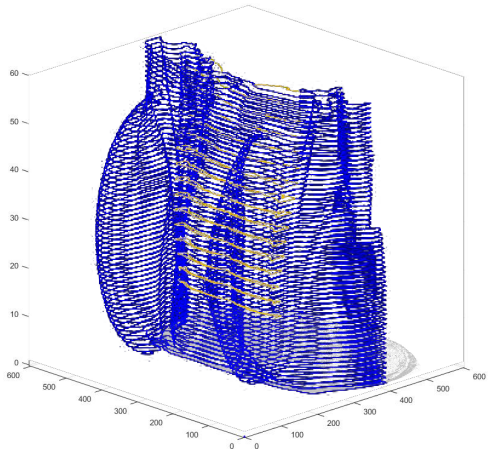


Figure 5.4: Breast Segmentation result for one of the patients

All the annotations were made using MARge Tool to appear in [Tei], whose interface is displayed in Figure 5.5.

5.1.2 Geometry Transformations

After the segmentation of the patient's MRI data, the meshes regarding the pre-surgery models were generated. While the MRI acquisition is done in prone, the tumor definition and posteriorly the surgical simulation need the model in a supine geometry. The transformation between the prone and the supine geometries were performed recurring to a Bio-mechanical simulator described in [VEH+16]. In order to perform these transformation a virtually state known as unloaded geometry was required. In the unloaded geometry, the force of gravity and other tension or stress forces are ignored [IAI+16]. The same biomechanical simulator is used to generate the pre-surgical model in a upright geometry (from the unloaded model) that will be further used. The

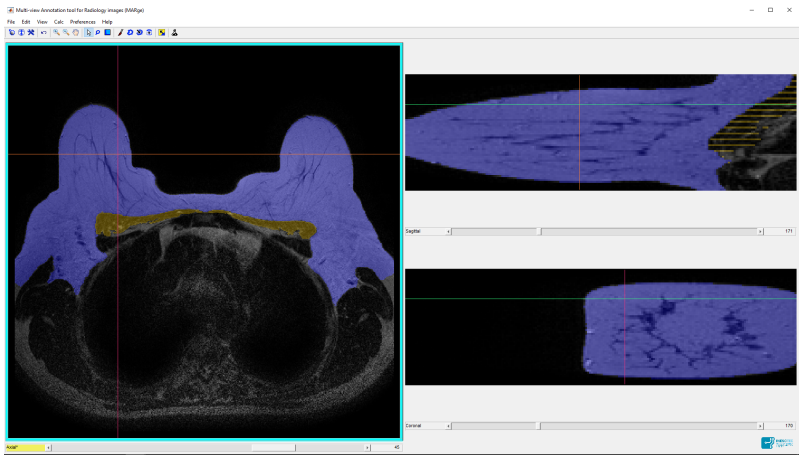


Figure 5.5: Interface of MARge tool used for the breast segmentation

aesthetic evaluation of the breast is performed in a upright position. The impact of changing the direction of gravity vector to a breast model is depicted in Figure 5.6.

5.1.3 Tumor's location definition

The tumor definition is done recurring to the tool described in chapter 4. The tool will take as input the pre-surgery model in a supine geometry and query the user for a tumor position. By selecting the tumor's position, it will be categorized according to one of the four quadrants of the breast also defined as regions, as shown in Figure 5.7:

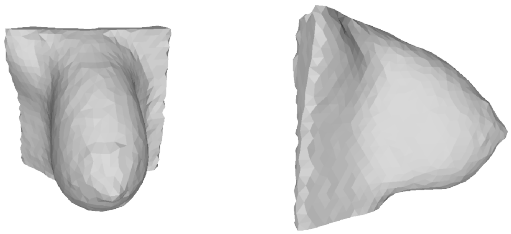
- R1 - Upper-Outer quadrant of the breast;
- R2 - Upper-Inner quadrant of the breast;
- R3 - Lower-Outer quadrant of the breast;
- R4 - Lower-Inner quadrant of the breast.

Thereafter a mesh containing the tumor in the defined region with one of the predefined sizes as follows will be generated:

- A small size - corresponding to 5% of the total breast's volume;
- A medium size - corresponding to 7.5% of the total breast's volume;
- And a large size - corresponding to 10% of the total breast's volume.

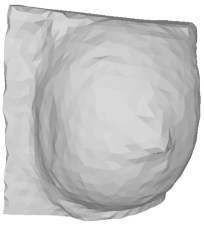
The percentages used as default for the tumor's size were obtained through discussion with physicians to understand the approximate size of the tumor in different stages of cancer detection.

The produced tumor will be represented by a cylinder centred on the chosen position for the tumor and perpendicular to the chest wall going from the pectoral muscle to the skin contour of



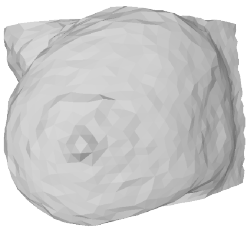
(a) 3D view of Prone geometry

(b) Lateral view of Prone geometry

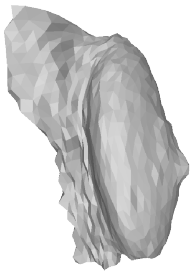


(c) 3D view of Unloaded geometry

(d) Lateral view of Unloaded geometry



(e) 3D view of Supine geometry



(f) Lateral view of Supine geometry

(g) 3D view of Upright geometry

(h) Lateral view of Upright geometry

Figure 5.6: 3D breast geometry transformations

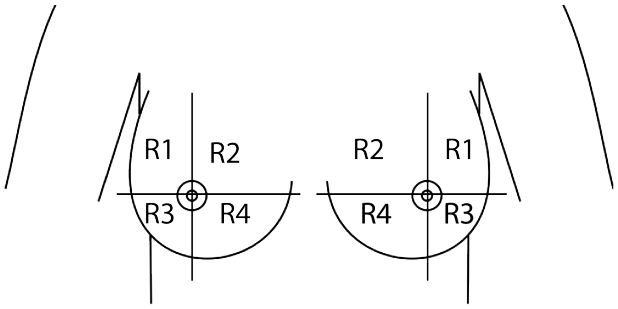


Figure 5.7: Breast quadrants used for tumor location

the breast. The cylinder's radius is calculated based on the tumor size. All the process is visually supervisioned and the results for the tumor need to be accepted by the user. As result a mesh with exact shape of the one used as input will be generated and posteriorly labelled to be used on the surgical simulation. The labelling process is explained in section 5.1.4.

5.1.4 Data Labelling

The labeling is done based on the mesh and what each point and element represents. There will be different labels for surface and volumetric information. The assigned labels indicate the following boundaries of the mesh: front, pectoral muscle or back, left and right or top and bottom. However, the lateral and top and bottom sides of the model are not deformed by the biomechanical neither the wound healing models. The volumetric information is labeled according with the material that it is represented and if it is considered healthy on the case of the breast or damaged on the case of the tumor. While the approach of [VEH+16] divides the volumetric elements in fat and glandular tissue, we assume the same type of material for the all breast. This material's properties are set in order to represent the breast's density according with its ACR (from I to IV). According to [EPJ+08], using two different material to represent both the fat and glandular tissues, instead of using only one material to represent the whole breast do not significantly affect the results.

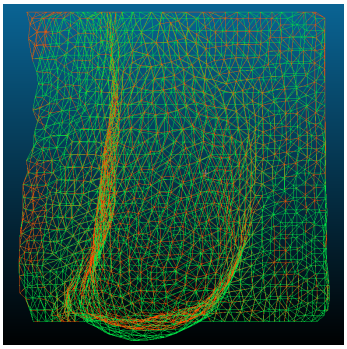
In order to visualize the models, a format file exchange is required. While this tool and the one presented on the next subsection require files in a *msh* format, to visualize the models the files must be parsed to a *ply* format. This is done through a parser developed in c++.

After defining damaged or healthy labels, the model will be transformed from supine into the unloaded geometry in order to serve as input on the next step, the surgical simulator, as described in subsection 5.1.2. All the information that is required on this transformations and for the wound healing application (including tumor location and volume and breast's density) is automatically generated by the tool presented in chapter 4.

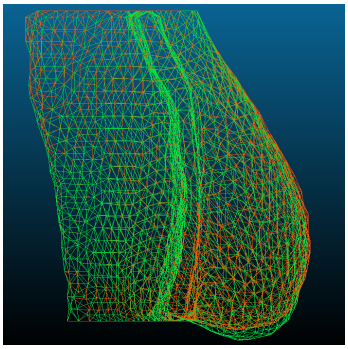
5.1.5 Wound healing simulation

The wound healing simulation, where the pos-surgical models are generated, is performed through the application of FEM and Multiscale Mechano-Biological expressions [VEH+16]. This will

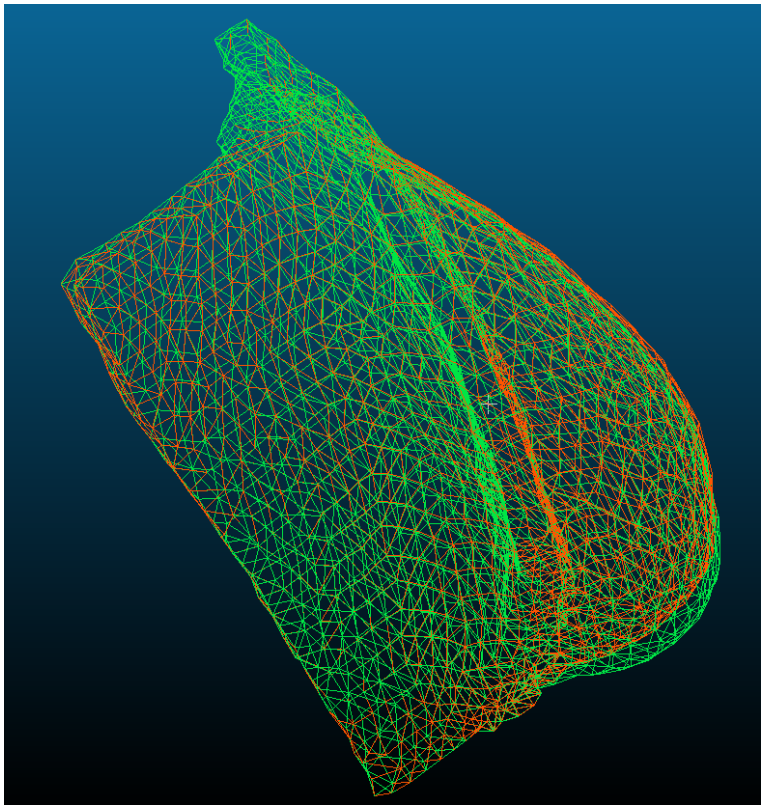
provide what would be the pos-surgical model of the patient's breast roughly 6 months (180 days) after the surgery, taking into consideration the patients breast density, the tumor's location and size that were artificially introduced before. Figure 5.8 demonstrates the breast generated by the wound healing simulation in an upright geometry.



(a) Frontal view



(b) Lateral view



(c) 3D view

Figure 5.8: Wound Healing simulation - comparison between the pre-surgical mesh (in green) and the pos-surgical mesh (in orange)

The final dataset currently provides pre and pos-surgical models for a total of 288 possible patients based on 6 initial real patients.

Despite of the annotation for both left and right breasts, using both breasts of the same patient would lead to very similar models to the natural symmetry of the human breasts. The initial real patients used to generate all the scenarios in the semi-synthetic dataset, were choose taking into account their breast volume, being classified as small, medium or large. For each patient regardless the breast size or laterality, models were generated for all the ACR (I to IV), for all the tumor regions (1 to 4) and for all the 3 sizes (1 to 3). These combinations are illustrated in Figure 5.9.

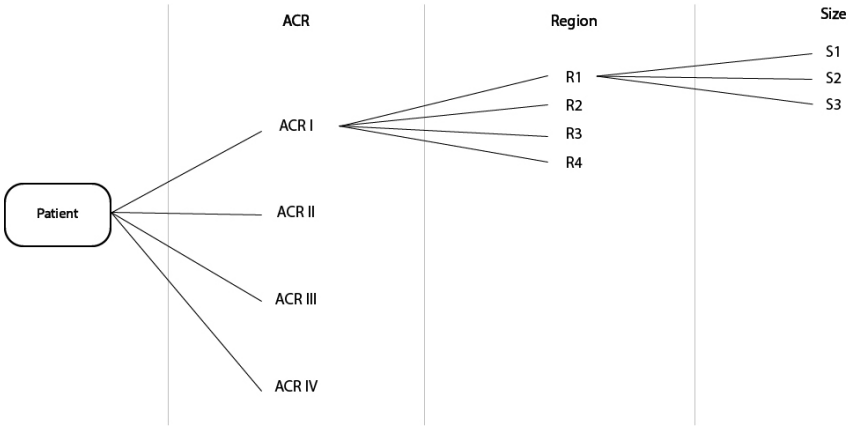


Figure 5.9: Representation of possible combination of clinical features for the dataset generation

All the dataset preparation steps (described previously in subsections 5.1.2, 5.1.3, 5.1.4 and 5.1.5) are illustrated in Figure 5.10.

5.2 Feature Analysis and Feature Construction

After the construction of the dataset, the data was analysed having into consideration the impact that the clinical features have on the healing simulation. Despite of the already existent and computed features, other features (like the distance between each point and the tumor's position) that were considered promising were computed and used in order to train the machine learning models.

5.2.1 Feature Analysis

The analysis of the clinical features was done by comparing the displacement of the corresponding points in the pre and pos-surgical models of the breast between variations of the same patient, where one of the clinical features was changing, and the others were kept the constant. The clinical features that were analysed were the tumor's size, the tumor's region and the ACR of the breast. With this study of the clinical features some conclusions were able to be drawn. The displacement of points resembles a "magnetic field" around the tumor's position and whether the

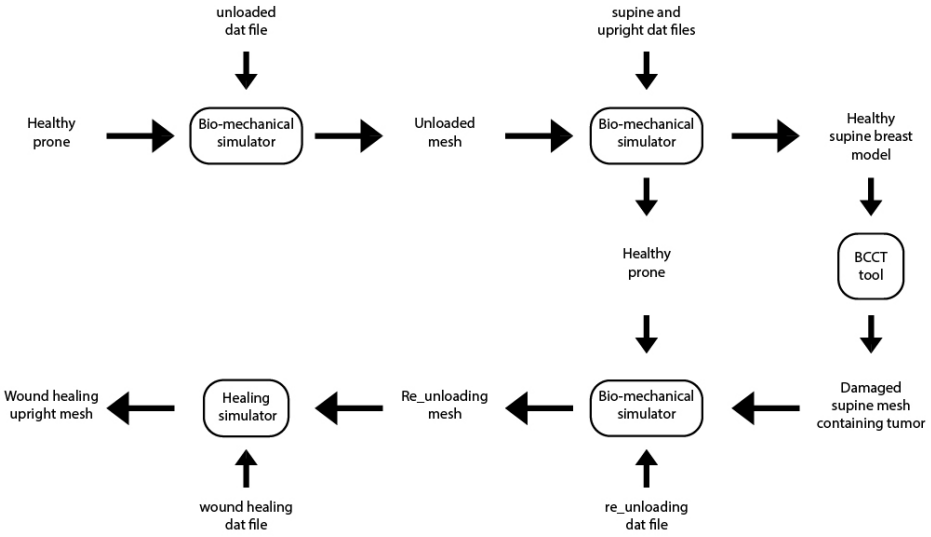


Figure 5.10: Workchart summarizing the breast's geometry transformations and wound healing simulation. The dat files are used for the Bio-mechanical simulator and the healing simulator containing the bio-mechanical and bio-chemical properties of the model's materials as well as the labelling definition.

tumor is located on an upper or lower region, the points of the lower region are always moved from their initial position, however, and as expected, when the tumor is located on a lower region, the displacement of the points will be larger and lead to a more profound deformation of the breast, as shown in Figure 5.11. The impact of the tumor size can also be noticed and as expected, a bigger tumor size leads to a greater impact of the breast deformation, represented in Figure 5.12. It is also possible to verify that the breast's density influence the deformations of the breast. A smaller ACR corresponding to a less dense breast will lead to a bigger displacement as represented in Figure 5.13.

Besides of these example, all the comparisons made and analyses in order to defer the impact of the clinical features on the breast deformation can be seen in appendix A.

5.2.2 Feature Construction

This analysis on clinical features also led to a few more conclusion about the deformation of the breast. Considering the new finding resultant from the feature analysis, some additional features were though to be helpful when trying to predict the new position of each point on the breast's point cloud. Despite of the findings regarding the clinical features, it was found is that the nearer surface points are to the tumour position, the more displacement they will suffer after the healing simulation. Having this in consideration, the euclidean distance and the difference for each axis, between the point itself and the tumor's center of mass, were computed and represented in both

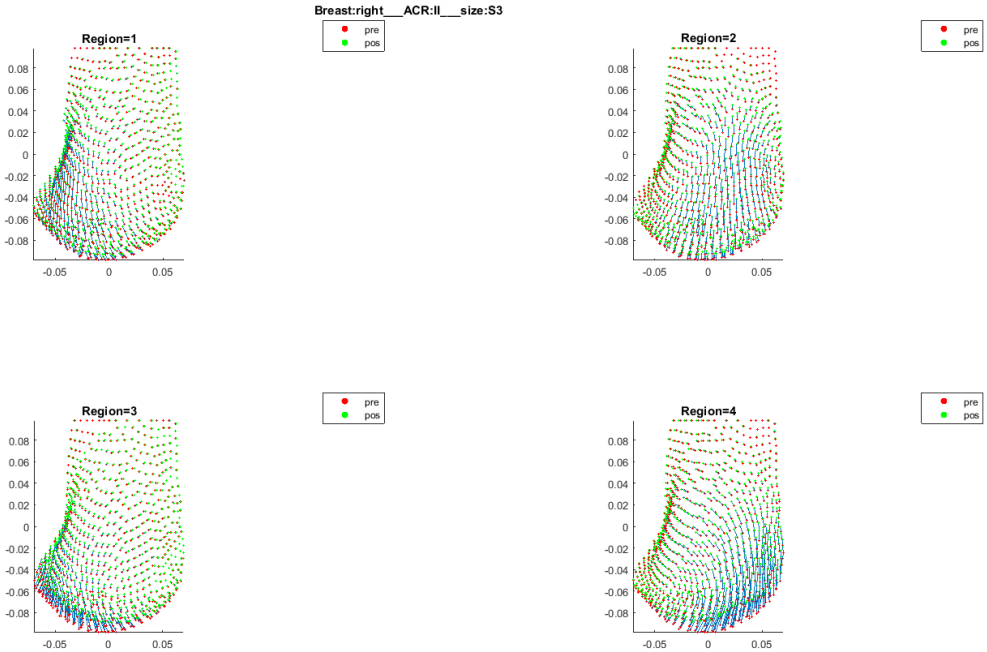


Figure 5.11: Example of the impact of tumor's location (region) on the breast's deformation

Cartesian and cylindrical coordinate. Being the cylindrical coordinates centred on the tumor's position.

5.3 Model Design

5.3.1 Machine Learning Models

As previously explained the intention of using ML (Machine Learning) Techniques is to predict the deformations caused by the BCS for the patients' breast. This prediction is carried out through the application of Regression Models which try to find a relation between the features training data, and the target values of the testing data. The target values will be the displacement of the points from the pre-surgical to the pos-surgical models of the breast.

5.3.1.1 Feature Representation

Despite of the tested scenario, the machine learning model will receive as input variables, the points of all the breasts in the training set. Each observation (as the input of the machine learning model) consists of a point from the pre-surgical model followed by several features regarding the characteristics of the breast. Initially, the training model will only have access to the points of the breast's surface, the front surface of the breast excluding the laterals, as seen in Figure 5.14. The

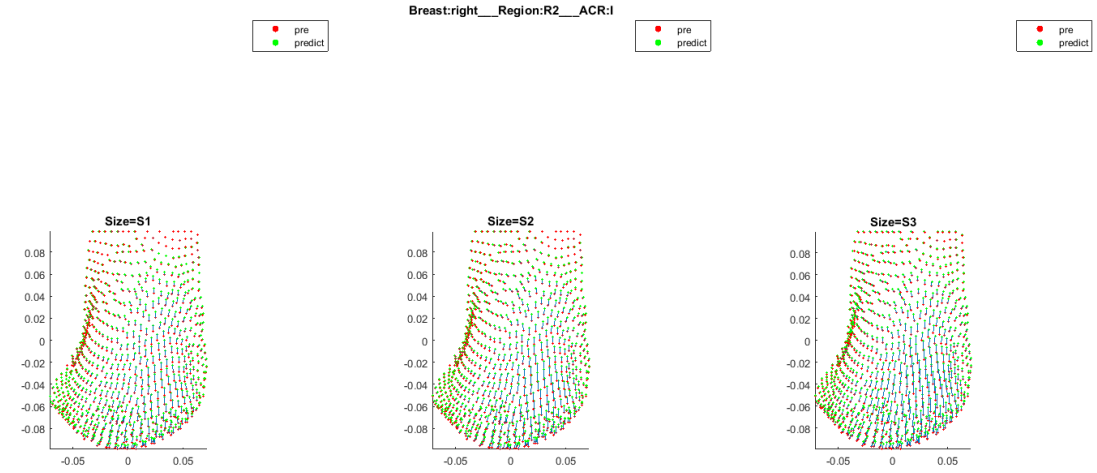


Figure 5.12: Example of the impact of tumor's size on the breast's deformation

models containing only the breast's surface that are going to be used contain an average of 629.33 point per model with a standard deviation of 106.68 points.

The used features, when categorical as the laterality, the ACR of the breast, the tumor's region and tumor's size, will be represented as "dummy variables" [1]. Other features such as the point coordinates, breast volume, and distances will be represented as real values.

The tumor's position defined by the breast's quadrant represents different quadrants of the breast according with the breasts laterality. For instance, while a region 4 (R4) on a left breast stands for the Lower left portion of the breast, the same region on the right breast refers to the Lower right portion. There are two possibilities in order to overcome this mismatch:

- Label the points according to the breast laterality, considering left and right points independent;
- Reflect one of the breast over yOz plane or vice-versa as represented in Figure 5.15;

Features like the breast laterality can be represented using categorical features or replaced by the breast's reflection. There are also features that will be represented in different ways, for instance the tumor's volume (real value of the breast's volume) and the tumor's size (categorical variable that represent the tumor as small, medium or large, regarding its size).

[1] http://imaging.mrc-cbu.cam.ac.uk/statswiki/FAQ/contint?action=AttachFile&do=get&target=int.pdf

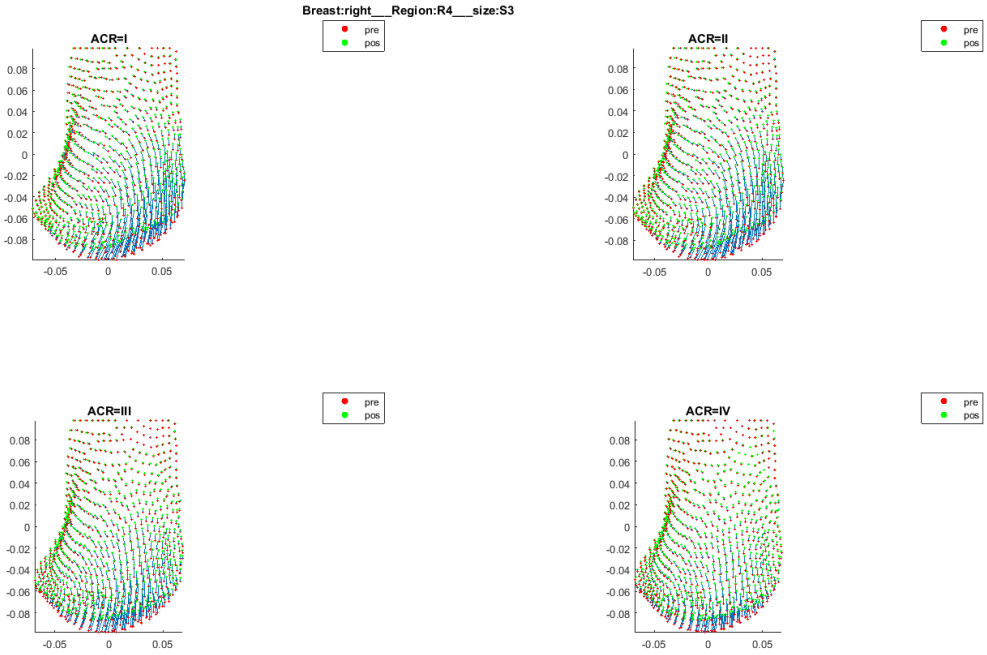


Figure 5.13: Example of the impact of breast density (ACR) on deformation

(a) Frontal view

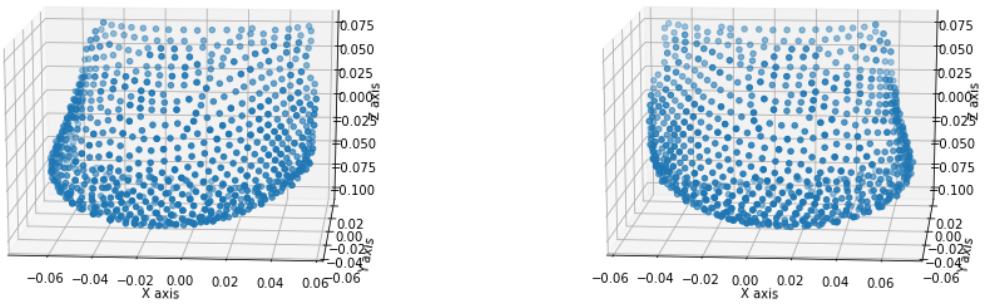
(b) Lateral view

Figure 5.14: Example of breast point cloud to be used as ML input

The Table 5.1 summarizes all the features that were used.

5.3.1.2 Labels

In order to obtain the shape of the breast after the BCS, the trained machine learning models predict the displacement of the points between the pre-surgical and pos-surgical models of the breast in each axis. This displacement is calculated by computing the difference in each axis



(a) left breast of a patient

(b) Reflection of the left breast of the patient

Figure 5.15: Example of breast point cloud's reflection

between the pos-surgical point of the 3D model and the correspondent point in the pre-surgical 3D model. Through the feature analysis described in subsection 5.2.1, it is evident that the axis where the points suffer a larger displacement is the z axis.

Having the dataset ready with both the features and labels computed, the dataset was split into training and testing sets. This division was done in two different ways:

- Splitting the dataset by applying a Leave-one-out (LOO) methodology. The dataset is divided into 6 sets of data (one per each initial real patient, containing 48 cases each), and using alternately each one of the sets as test set and the remaining 5 sets as training set;
- Randomly split the 288 cases of the complete dataset using 1/6 of the cases as testing set and the remaining cases (5/6) as training set.

Regarding the validation set, the models use 10-fold cross validation. As expected and as can be seen in chapter 6, randomly splitting the data leads to much better although misleading results than the other approach for the splitting of the dataset into training and testing sets. This is justified by the similarity of the breast within the training and the testing set, since both of them are simulations of the same initial patient despite of the different clinical features. The use of this biased testing set allowed the representation of a best case scenario for the prediction of the breast's deformation.

5.3.1.3 Machine Learning Algorithms

In order to train the machine learning models that will be used to predict the displacement of the pre-surgical model to the pos-surgical, the following machine learning algorithms were used in order to solve this regression problem.

- Multilayer Perceptron (MLP)

MLP, also known as feedforward neural networks (FFNN) is a type of Artificial Neural Networks widely used when targetting clinical medical issues [FFA+00]. They consist of a set of nodes connected by edges that simulate the behaviour of neurons of the human brain. [NJ16]

The node is responsible to compute the weighted sum of the inputs $z = \sum_{j=1}^{m} w_j x_j$ and to apply an activation function $y = \varphi(z+b)$, where x_j is the input for input link j , w_j the weight for the same input, y the output of the neuron, b the optimal bias parameters added to the input and φ is an activation function.

An example of a MLP is represented in Figure 5.16 consisting of multiple layers of fully connected neurons. The represented layers differ in three types:

- Input layers - only transport all the inputs to the next layer;
- Hidden layers - adjust the weighed sum of the inputs, compute the activation function and output the values to the next layer;
- Output layers - perform the same computations that hidden layers do and output the values as the network's result.

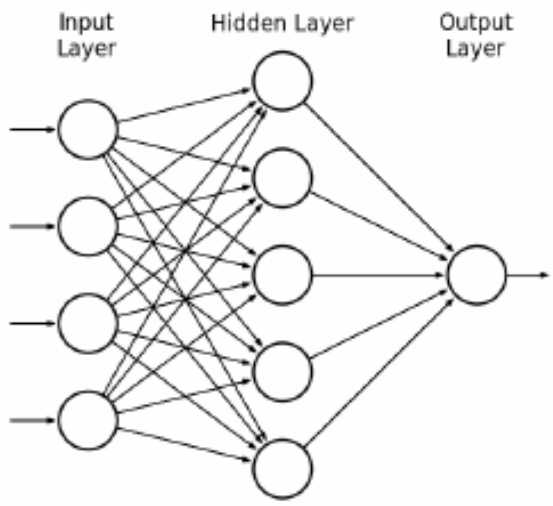


Figure 5.16: Representation of a MLP [OoBIBP+16]

The number of neurons highly depends on the amount and properties of the training data, as well as the number of hidden layers. And the Learning rate impacts the step used to adjust the inputs' weight.

- Random Forest (RF)

  The Random Forest (RF) algorithm is based on using multiple decisions trees, mitigating the negative aspects of decision trees by ignoring some of the input properties and increasing the performance of the algorithm. Each decision tree on the RF system, only consider a

random subset of the input data, and consider a limited number of features smaller than the number of total features. The output of the several decision trees is then averaged and used as output of the RF [NJ16].

The splitting criteria for regression problems in RF algorithms to divide the root or leaf into more leafs is calculate through $RSS = \sum_{LEFT}(Y_i - Y_L^*)^2 + \sum_{RIGHT}(Y_i - Y_R^*)^2$, where RSS stands for Residual sum of the squares, Y_i stands for the current node and Y_L^* and Y_R^* stands for the mean value of y for both the left and right nodes, respectively [Cut13].

Figure 5.17 presents an example of a decision tree regression result.

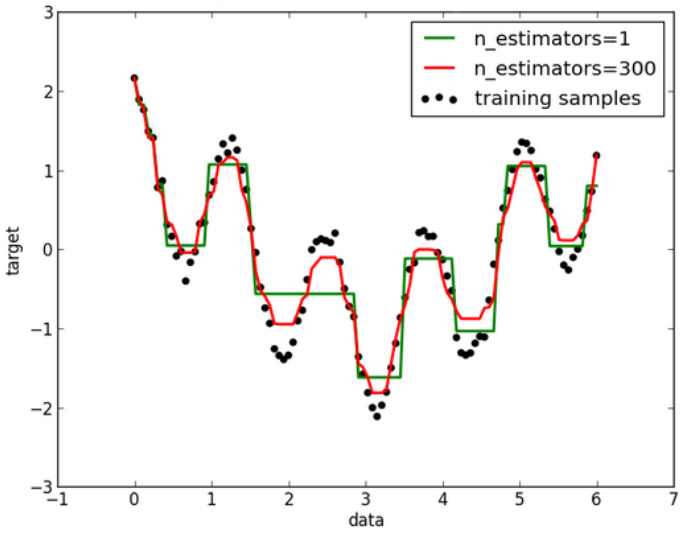


Figure 5.17: Example of a RF result. Comparison of results of a RF with only a decision tree (green) and a RF with 300 decision trees (red). The scatter point represent the data used to train the RF. [2]

This type of algorithm has as great advantage not requiring a prior feature section, and despite of being promising, it is only capable of predicting the regression of a variable. In this case, in order to predict the total displacement of the point between models, three individual training models are required, one for each Cartesian axis.

- Multioutput Regression (MOR)

  Multioutput regression also known as multi-variate or multi-target regression, has been used on multiple applications across several fields of study [BVBM]. This ML approach allows the prediction of several output variables for each entry of the model. [3] Other advantage of this type of approach is that the produced models are, generally, simpler and more computational efficient.

[2] https://www.quora.com/How-does-random-forest-work-for-regression-1
[3] http://scikit-learn.org/stable/modules/multiclass.html#multiclass

One of the several multioutput regression method consists on the adaptation of other ML algorithms such as RF in order to allow them to predict several output variables. This type of multioutput regressors requires the identification of the dependencies among the target variables. When compared to regular RF, multi-target regression trees usually require a smaller number of trees for all the variables, and allow a more informed understanding of the dependencies among the several target variables of the model [BVBM].

5.3.1.4 Implementation

As already stated, the presented algorithms were implemented aiming to predict the displacement between the pre-surgical and the pos-surgical models. The displacement for x axis, y axis and z axis are respectively represented as ∂x, ∂y and ∂z.

The 3 previously described machine learning algorithms were implemented using 10 fold cross validation in order to automatically tune each algorithm parameters leading the algorithm to achieve the minimum root mean squared error of the model. The 3 algorithms were used to train both of the splits of the dataset (LOO and the random split using 1/6 of the data as the test set).

A short feature selection was done to understand which features should be used as well as their representations. The feature selection was done recurring to recursive feature elimination [4], and then a trial was done using MLP. Regarding the training of RF and MOR, this study was not necessary since both algorithms support a built-in feature selection method that measures the feature's importance [5]. The last two algorithms led to the training of random forest models for each axis (x, y and z axis); and the training of multi output regressor models, considering the ∂z, ∂y, ∂x as labels; ∂y and ∂x as labels; and ∂x and ∂y as labels.

In order to implement the random forests and multi-layer perceptron algorithms a R package named caret [6] was used. Multi-output regressor was implemented with a R package called randomForestSRC package [7].

5.3.2 Naive Model

In order to understand the impact of ML techniques to predict the deformation, a naive method was created to generate the shape of the breast using common sense and conclusion arising from the feature analysis presented in section 5.2.1.

Two alternatives for the naive method were developed, where the displacement between the pre-surgical and pos-surgical models of the breast were computed as follows:

1. The first alternative consisted on calculating the average displacement of the points on each geometric quadrant of the breast. The average displacement of each quadrant is calculated

[4] https://topepo.github.io/caret/recursive-feature-elimination.html
[5] https://topepo.github.io/caret/feature-selection-overview.html
[6] https://cran.r-project.org/web/packages/caret/caret.pdf
[7] https://cran.r-project.org/web/packages/randomForestSRC/randomForestSRC.pdf

based on the mean displacement of the points on the same quadrant in similar breasts. These similar breast concerns all the breasts on the dataset, that were generated from a different initial patient, with the same breast properties: breast's laterality and density and the same tumor's properties.

For instance, if the breast that is under the naive method has the tumor located in region 2, the average displacement for the correspondent quadrant will be calculated based on the other patients with the tumor on the same region, the same tumor's size and the same breast's laterality and density. For the regions were the tumor is not located, the average displacement will be calculated based on the breasts of the dataset, that do not have the tumor located on that region and with the same breast's properties and tumor's size. Therefore, the quadrants are defined based on the geometric center of the breast.

In spite of calculating an average displacement for the all quadrants of the breast, only a portion of that displacement is applied to the pre-surgical model. Each quadrant was divided into 3 stacks across the z axis, and a portion of the total displacement was applied based on how much the points on that region would normally move. The stack division is done equally, however the portion of the applied displacement on the x, y and z axis differs according to the quadrants (lower or upper quadrant).

2. The second approach of the naive method, is divided into two steps. In the first step, and similar with what happens on the first approach, the breast is divided into quadrants, however considering the nipple's position but not the geometric center of the breast. In the second step, and taking into consideration all breasts in the dataset, the average displacement of points is calculated for each quadrant of the breast. When considering the quadrant where the tumor is located, the points below the tumor are updated by subtracting their height from the mean displacement previously calculated; When considering the not affected quadrants (without the tumor located on them), the points' position is updated. Calculating the ratio between the average displacement of the non-affected quadrants and the distance from the point to the tumor's center of mass, will give a value that, when subtracted from the point's previous position will result on the new position of the point. The computed displacements are also multiplied by the breasts' ACR using a factor of 1.2, due to the relation between the breast's ACR and the breast's deformation.

5.4 Results Validation

Despite of the use of cross validation and the parameter tuning used on the implementation of the machine learning models that tried to decrease the root mean squared error between the predicted data and the given labels, the obtained results were evaluated using some evaluation metrics. As previously described, the given machine learning models had as goal to predict the displacement of each point in the breast's 3D model in one or more axis. Summing the predicted values to

the correspondent point of the pre-surgical breast's model, ideally would produce the pos-surgical model of the breast.

5.4.1 Evaluation Metrics

There were considered two types of evaluation metrics:

- Visual metrics;

  The visual metrics consists of visual comparison between pre-surgical, pos-surgical and predicted models of the breast. This evaluation allows to easily gather some conclusion regarding the models performance.

- Distance metrics.

  The distance metrics allows to obtain a more accurate perception of the model's performance.

5.4.1.1 Distance metrics

The distances computed for evaluating the models consider both local and global distances: while, the local distance, compare the models point by point, the global distances is measured in both directions comparing each point of a model to another breast's model and a breast's model to a point of another model. Regarding the local metric, the distances to be measured are the one between the predicted model and the pos-surgical model (*predicted to pos*); the distance between the pre-surgical to the pos-surgical models (*pre to pos*); and the distance between the predicted model and the pre-surgical model (*predicted to pre*). Regarding the global metric, the same distances will be measures, however in both directions: *predicted to pos*; *pos to predicted*; *pre to pos*; *pos to pre*; *predicted to pre*; and *pre to predicted*.

On either case the following values are computed:

- Mean of Euclidean distance;
- Standard Deviation of the Euclidean distance;
- Hausdorff distance (maximum of the euclidean distances).

These values are calculated considering the 3 coordinates of each point (3D), or only the point's projection in one of the Cartesian coordinate system axis(1D).

5.5 Summary

The present chapter, described the complete process that was followed in order to predict the deformation caused by BCS. Initially the dataset preparation was described, and some additional information was given, in order to understand its construction. Then a analysis of the clinical features was also detailed and results presented in the chapter. Thereafter the models used in order to

predict the breast deformations were explained. Regarding the introduced machine learning techniques, a short background of the regression models was provided and all the relevant information about their implementation and application is outlined. At last, the evaluation metrics were also described.

All the results and further conclusion arising from the detailed methodology are explored in chapter 6.

Feature	**Feature Description**	**Variable name**	**Variable Description**
Point coordinates	Point's representation in a Cartesian coordinate system	x_coord	coordinate of point in x axis
		y_coord	coordinate of point in y axis
		z_coord	coordinate of point in z axis
Distance's projection between point and tumor's center of mass	Distrance's projection between the position of the breast's point and the tumor's center of mass. This difference can be represented in a Cartesian coordinate system (variables x_dist, y_dist and z_dist) or through cylindrical coordinates (variables theta, rho and pol_z)	x_diff	distance's projection between the position of the point and the tumor's center of mass in a Cartesian coordinate system (x axis)
		y_diff	distance's projection between the position of the point and the tumor's center of mass in a Cartesian coordinate system (y axis)
		z_diff	distance's projection between the position of the point and the tumor's center of mass in a Cartesian coordinate system (z axis)
		theta	distance's projection between the position of the point and the tumor's center of mass in a cylindrical coordinate system (theta angle)
		rho	distance's projection between the position of the point and the tumor's center of mass in a cylindrical coordinate system (rho angle)
		pol_z	distance's projection between the position of the point and the tumor's center of mass in a cylindrical coordinate system (z value)
Distance	Euclidean distance between the point and the tumor's center	dist_Tpt	Euclidean distance between the point and the tumor's center of mass
Breast's Volume	Pre-surgical breast's volume	b_vol	Volume of the pre-surgical breast's model
Tumor's Size and Volume	Representation of the tumor's volume, either using the real value or categorical variables	t_vol	Real value of the tumor's volume
		t_size_a	Categorical variable to represent the tumor's size (small)
		t_size_b	Categorical variable to represent the tumor's size (medium)
		t_size_c	Categorical variable to represent the tumor's size (large)
Breast's Laterality	Representation of the breast's laterality (left or right breast)	lat_a	Categorical variable to represent the breast's laterality (Right Breast)
		lat_b	Categorical variable to represent the breast's laterality (Left Breast)
Breast's Density (ACR)	Representation of the breast's density using ACR	acr_a	Categorical variable to represent the breast's density (ACR I)
		acr_b	Categorical variable to represent the breast's density (ACR II)
		acr_c	Categorical variable to represent the breast's density (ACR III)
		acr_d	Categorical variable to represent the breast's density (ACR IV)
Tumor's Location (Region)	Representation of the tumor's location regarding the breast's quadrant	reg_a	Categorical variable to represent the tumor's located in R1
		reg_b	Categorical variable to represent the tumor's located in R2
		reg_c	Categorical variable to represent the tumor's located in R3
		reg_d	Categorical variable to represent the tumor's located in R4

Table 5.1: List of Features

Chapter 6

Results and Discussion

This chapter intends to present the results for both the several performed machine learning runs using different algorithms and different features as well as different combinations of features and the naive implementations. By analysing the results using visual cues and the evaluation metrics previously defined in chapter 5, the upcoming findings suggest if the machine learning techniques are correctly predicting the breast deformation. The results will also be further analysed in order to understand on which regions of the breast, the prediction is more inaccurate in terms of distance between the predicted and the real data.

The current chapter is divided into 3 sections, being the first two, section 6.1 and section 6.2, used to shown the visual and evaluation metrics' results for the naive methods and the machine learning models respectively. Section 6.3 will present the regions of the breast where the predictions are more inaccurate. All the results and intermediary conclusions will be used to draw the final conclusions present in chapter 7.

6.1 Naive Method Results

As described before, two different naive methods approaches were tried in order to predicted the deformation caused by BCS. The implementation of this approach has as goal, understanding the improvement provided by the use of machine learning techniques instead of using methods based on common sense and the findings resultant from the feature analysis in section 5.2.1.

Table 6.1 shows the global evaluation metrics for the the naive method described in section 5.3.2 that uses the geometric center of the breast in order to divide it into quadrants. Table 6.2 represents the local evaluation metrics in the same conditions. As shown the prediction made by the naive method leads to the movement, not very significant, of the breast's. This effect can be visualized in Figure 6.1. Using this method leads to an over prediction of the displacement in cases similar to the one represented in Figure 6.1a, where the the breast suffers a minor deformation

caused by the BCS and to an under prediction of the displacement in cases similar to the one represented in Figure 6.1b, where the breast's deformation caused by the BCS is more severe.

			predicted to pos	pos to predicted	pre to pos	pos to pre	predicted to pre	pre to predicted
3D	**Euclidean Distance**	**Mean**	1.731	1.725	1.758	1.731	1.420	1.421
		Standard Deviation	1.130	1.113	1.333	1.277	1.064	1.394
	Hausdorff Distance		5.563	5.539	6.513	6.317	2.939	6.452
1D	**Euclidean Distance**	**Mean**	0.141	0.123	0.160	0.110	0.113	0.083
		Standard Deviation	0.210	0.145	0.346	0.110	0.142	0.243
	Hausdorff Distance		1.631	1.203	3.168	0.704	0.787	3.325

Table 6.1: Global Evaluation Metrics for the first approach of the naive method

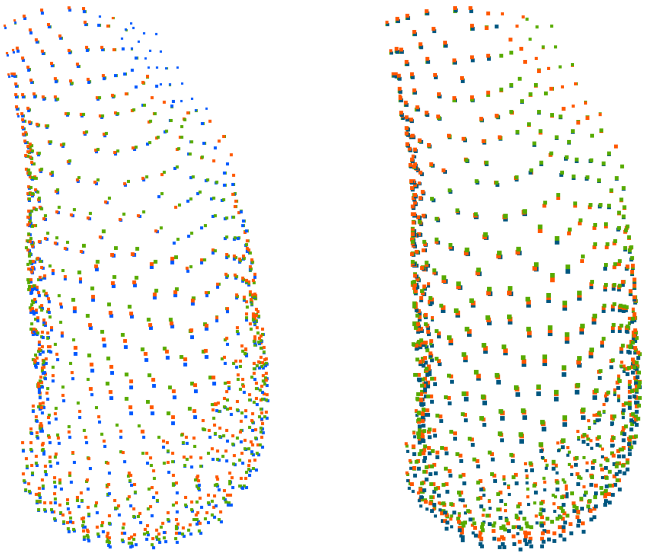
			predicted to pos	pre to pos	predicted to pre
3D	**Euclidean Distance**	**Mean**	1.980	2.206	1.426
		Standard Deviation	1.503	1.920	1.070
	Hausdorff Distance		7.000	8.410	2.939
1D	**Euclidean Distance**	**Mean**	1.601	1.904	1.397
		Standard Deviation	1.530	1.929	1.085
	Hausdorff Distance		6.635	8.092	2.926

Table 6.2: Local Evaluation Metrics for the first approach of the naive method

The global and local evaluation metrics for the second implementation of the naive method also described in section 5.3.2 are represented respectively in Table 6.3 and Table 6.4. Despite of the improvement on the evaluation metrics, the same effect represent in Figure 6.1 still occurs as shown in Figure 6.2.

			predicted to pos	pos to predicted	pre to pos	pos to pre	predicted to pre	pre to predicted
3D	**Euclidean Distance**	**Mean**	1.523	1.503	1.758	1.731	1.414	1.415
		Standard Deviation	1.194	1.129	1.333	1.277	1.067	1.069
	Hausdorff Distance		5.327	5.137	6.513	6.317	2.939	2.939
1D	**Euclidean Distance**	**Mean**	0.127	0.113	0.160	0.110	0.073	0.090
		Standard Deviation	0.194	0.120	0.346	0.110	0.101	0.193
	Hausdorff Distance		1.685	0.851	3.168	0.704	0.618	2.091

Table 6.3: Global Evaluation Metrics for the second approach of the naive method



(a) Example of a breast with a less deformation caused by BCS

(b) Example of a breast with a more noticeable deformation caused by BCS

Figure 6.1: Comparison between pre-surgical, pos-surgical and predicted through a naive method breast's models. The pre-surgical model is displayed in blue; the pos-surgical models displayed in green, and the naive predicted model displayed in orange.

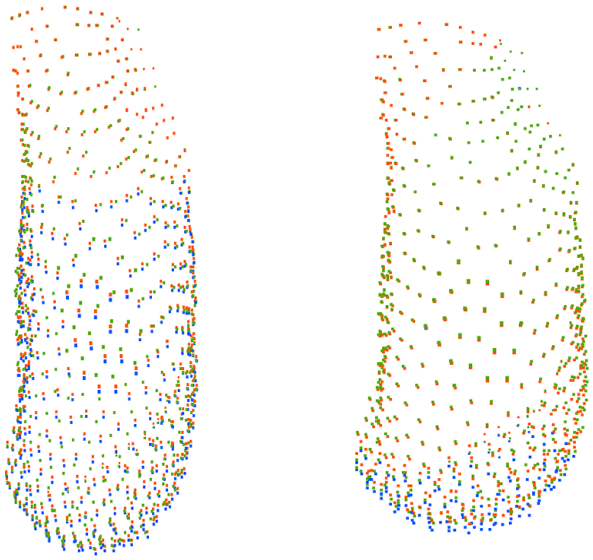
			predicted to pos	**pre to pos**	**predicted to pre**
3D	**Euclidean Distance**	**Mean**	1.634	2.206	1.723
		Standard Deviation	1.497	1.920	1.124
	Hausdorff Distance		5.763	8.410	3.152
1D	**Euclidean Distance**	**Mean**	1.231	1.904	1.397
		Standard Deviation	1.055	1.929	0.924
	Hausdorff Distance		5.174	8.092	5.191

Table 6.4: Local Evaluation Metrics for the second approach of the naive method

6.2 Machine Learning Results

The different algorithms described in section 5.3.1.3 were used to train the several prediction models.

The following sections present the results for each ML algorithm in order to understand what led to a better prediction, as well as the features and labels that should be used on each case.



(a) Example of a breast with a lesser deformation caused by BCS

(b) Example of a breast with a more noticeable deformation caused by BCS

Figure 6.2: Comparison between pre-surgical, pos-surgical and predicted through a naive method breast's models. The pre-surgical model is displayed in blue; the pos-surgical model is displayed in green, and the naive predicted model is displayed in orange.

6.2.1 Random Forest Results

In order to predict the displacement of the breast's surface points using RF, three different models were built: one for each axis. Despite of the 3 models, the evaluation was done considering the simultaneous outcome of the three models.

On an initial approach, only considering the surface points in order to train the models, features such the points' coordinates, the distance from the point to the tumor's center of mass, the breast's volume, the tumor's volume and the remaining clinical features as dummy variables. as the points were considered. Another trials were carried out, where the features were represented differently or even omitted. On one of those trials, instead of using the breast's laterality as a categorical variable, the right breast were reflected and considered left breast, leading to worse results in terms of a more inaccurate displacement prediction.

Both global and local evaluation metrics are respectively displayed in Table 6.5 and Table 6.6 for the LOO train/test splitting and in Table 6.7 and Table 6.8 for the biased train/test split. The biased split (as described in section 5.3.1.2) was constructed by randomly splitting the dataset cases into train and test sets. Since all the cases in the dataset were generated from 6 initial real patients, there is a great probability of a case in the test set having a very similar cases on the training set. By using this is possible to represent an ideal situation of the breast's deformation prediction and understand the results of the best case scenario of the prediction model.

In order to train the models, whose results were previously presented, were trained using an

			predicted to pos	pos to predicted	pre to pos	pos to pre	predicted to pre	pre to predicted
3D	**Euclidean Distance**	**Mean**	1.103	1.102	1.758	1.731	1.848	1.866
		Standard Deviation	0.811	0.810	1.333	1.277	1.228	1.266
	Hausdorff Distance		4.062	4.031	6.513	6.317	5.580	5.738
1D	**Euclidean Distance**	**Mean**	0.105	0.104	0.160	0.110	0.119	0.158
		Standard Deviation	0.117	0.113	0.346	0.110	0.116	0.313
	Hausdorff Distance		0.970	0.875	3.168	0.704	0.714	3.049

Table 6.5: Global Evaluation Metrics for RF models using LOO train/test split

			predicted to pos	pre to pos	predicted to pre
3D	**Euclidean Distance**	**Mean**	1.151	2.206	2.190
		Standard Deviation	0.899	1.920	1.643
	Hausdorff Distance		4.464	8.410	6.502
1D	**Euclidean Distance**	**Mean**	0.909	1.904	1.979
		Standard Deviation	0.931	1.929	1.578
	Hausdorff Distance		4.383	8.092	5.907

Table 6.6: Local Evaluation Metrics for RF models using LOO train/test split

			predicted to pos	pos to predicted	pre to pos	pos to pre	predicted to pre	pre to predicted
3D	**Euclidean Distance**	**Mean**	0.674	0.672	1.799	1.773	1.772	1.784
		Standard Deviation	0.600	0.597	1.379	1.325	1.236	1.262
	Hausdorff Distance		3.199	3.127	6.550	6.326	5.398	5.539
1D	**Euclidean Distance**	**Mean**	0.096	0.096	0.162	0.111	0.119	0.160
		Standard Deviation	0.103	0.102	0.349	0.111	0.116	0.328
	Hausdorff Distance		0.781	0.761	3.274	0.718	0.725	3.238

Table 6.7: Global Evaluation Metrics for RF models using the train/test biased split

automatic model tuning [1], that tries to find the best parametrization of the model to the problem. In the presented case, the model parameters were as follows: $mtry = 8$; $n\_trees = 250$; $node\_size = 1$. *n_trees* represents the number of decision trees used for training the model, *mtry* represents the number of variables sampled at each split, and *node_size* represents the minimum size of terminal

[1] http://machinelearningmastery.com/tuning-machine-learning-models-using-the-caret-r-package/

			predicted to pos	pre to pos	predicted to pre
3D	**Euclidean Distance**	**Mean**	0.680	2.196	2.096
		Standard Deviation	0.615	1.917	1.636
	Hausdorff Distance		3.256	8.337	6.448
1D	**Euclidean Distance**	**Mean**	0.643	2.115	2.032
		Standard Deviation	0.730	2.115	1.719
	Hausdorff Distance		3.777	8.728	6.483

Table 6.8: Local Evaluation Metrics for RF models using the train/test biased split

nodes on the model. [2]. The feature importance computed by these models is also represented in Table 6.9 and the visual results are presented in Figure 6.3.

Features	∂**x**	∂**y**	∂**z**
x_coord	48.10	55.22	76.37
y_coord	74.93	101.35	78.38
z_coord	85.16	71.69	35.85
x_diff	82.15	60.54	85.77
y_diff	56.29	71.69	61.87
z_diff	73.54	127.16	59.96
dist	57.78	58.70	45.61
b_vol	25.08	40.52	59.92
t_vol	103.19	132.59	168.23
lat_a	18.07	12.85	26.63
lat_b	17.41	13.38	26.88
acr_a	25.58	27.89	27.64
acr_b	28.95	23.58	20.25
acr_c	30.48	23.95	24.78
acr_d	30.69	19.70	22.73
reg_a	27.22	21.95	12.76
reg_b	28.24	28.29	22.37
reg_c	29.95	23.56	17.61
reg_d	41.45	22.98	27.28

Table 6.9: RF feature importance

Despite of the previous scenario only considered the surface points of the patient's breast models, using additional points such as internal points, can lead to a better prediction of the deformation. The internal points were only used to train the model. The global and local distance metrics for this scenario are respectively shown in Table 6.10 and Table 6.11. These results were

[2] https://www.rdocumentation.org/packages/randomForest/versions/4.6-12/topics/randomForest

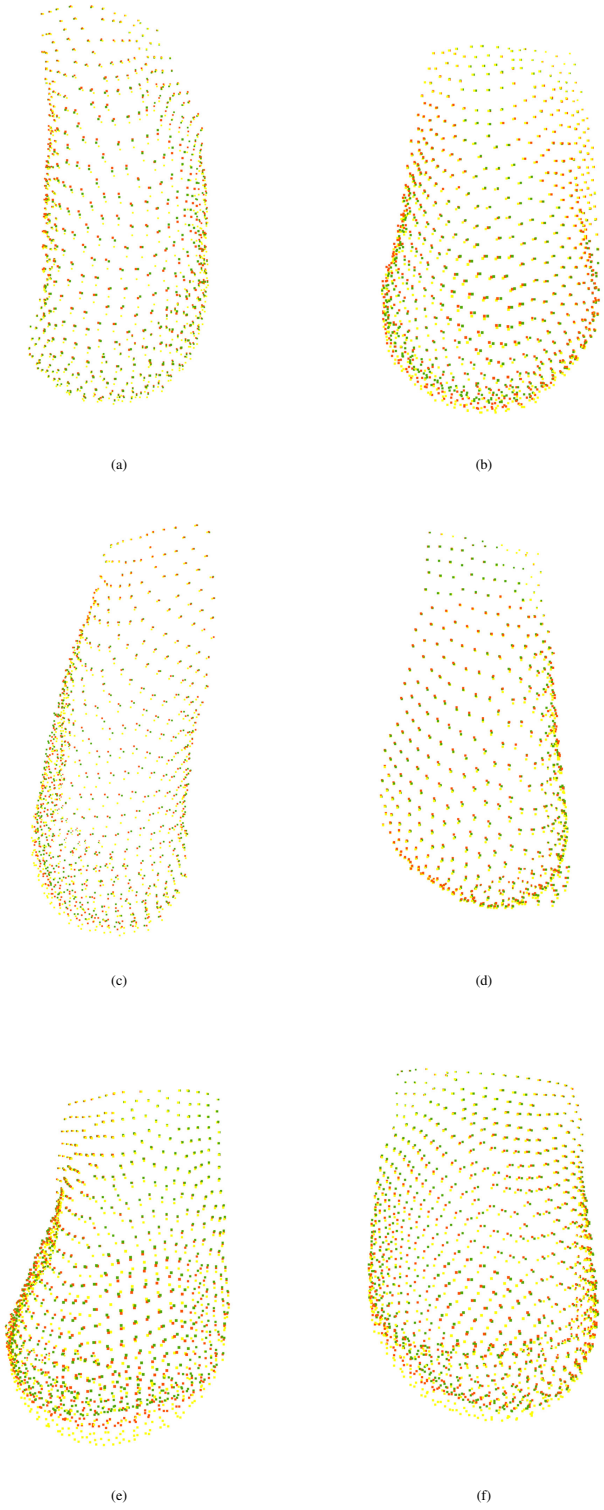


Figure 6.3: Visual examples of the prediction results obtained by the RF prediction models. The pre-surgical model is displayed in yellow; the pos-surgical model is displayed in green; The predicted model is displayed in red.

achieved by using the LOO split and the same model parameters. The evaluation metrics regarding the results of the models trained with both surface and internal points, with the biased split are described in Table 6.12 and Table 6.13.

			predicted to pos	pos to predicted	pre to pos	pos to pre	predicted to pre	pre to predicted
3D	**Euclidean Distance**	**Mean**	1.044	1.043	1.555	1.529	1.652	1.670
		Standard Deviation	0.840	0.837	1.382	1.319	1.257	1.296
	Hausdorff Distance		4.875	4.817	7.107	6.406	5.533	5.923
1D	**Euclidean Distance**	**Mean**	0.055	0.055	0.080	0.055	0.060	0.083
		Standard Deviation	0.063	0.068	0.236	0.060	0.062	0.238
	Hausdorff Distance		0.659	0.789	3.154	0.477	0.465	3.287

Table 6.10: Global Evaluation Metrics for RF models considering breast internal points with LOO train/test split

			predicted to pos	pre to pos	predicted to pre
3D	**Euclidean Distance**	**Mean**	1.077	1.919	1.912
		Standard Deviation	0.911	1.940	1.611
	Hausdorff Distance		5.352	9.510	6.620
1D	**Euclidean Distance**	**Mean**	0.817	1.693	1.759
		Standard Deviation	0.889	1.908	1.620
	Hausdorff Distance		5.003	9.147	6.475

Table 6.11: Local Evaluation Metrics for RF models considering breast internal points with LOO train/test split

			predicted to pos	pos to predicted	pre to pos	pos to pre	predicted to pre	pre to predicted
3D	**Euclidean Distance**	**Mean**	0.639	0.638	1.592	1.562	1.653	1.679
		Standard Deviation	0.614	0.610	1.391	1.318	1.336	1.394
	Hausdorff Distance		3.979	3.932	7.054	6.252	5.957	6.452
1D	**Euclidean Distance**	**Mean**	0.050	0.049	0.088	0.056	0.058	0.083
		Standard Deviation	0.062	0.053	0.284	0.060	0.062	0.243
	Hausdorff Distance		0.773	0.480	3.695	0.461	0.471	3.325

Table 6.12: Global Evaluation Metrics for RF models considering breast internal points with random train/test split

			predicted to pos	pre to pos	predicted to pre
3D	**Euclidean Distance**	**Mean**	0.644	2.019	2.001
		Standard Deviation	0.630	2.046	1.826
	Hausdorff Distance		4.090	9.927	7.420
1D	**Euclidean Distance**	**Mean**	0.501	1.732	1.631
		Standard Deviation	0.624	1.941	1.648
	Hausdorff Distance		3.870	9.190	6.710

Table 6.13: Local Evaluation Metrics for RF models considering breast internal points with random train/test split

By comparing the results of the two distinct scenarios previously analysed, it is possible to conclude that considering the internal points of the breast's model leads to better results.

6.2.2 Multilayer Perceptron Results

In order to try achieving better results other ML algorithms such as MLP were tested. However, unlike RF, other ML algorithms require a prior feature selection in order to achieve reliable values. In case of MLP, a Recursive Feature Elimination [3] (RFE) technique was used in order to understand the features to be used on the model training.

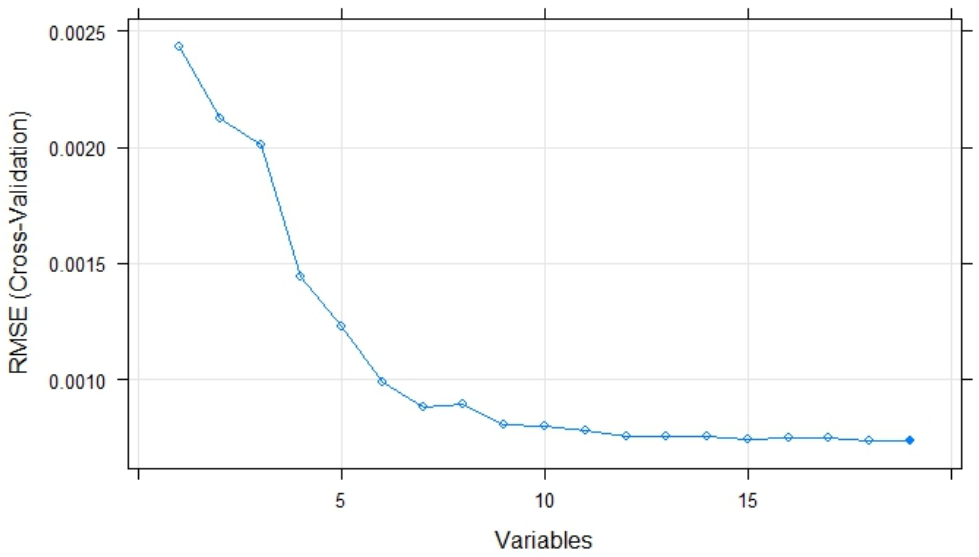


Figure 6.4: Recursive Feature Elimination of variables used for MLP

[3] https://topepo.github.io/caret/recursive-feature-elimination.html

Considering the result of the feature selection, it is possible to conclude that all the features lead to a decrease of the root mean squared error (RMSE), this way, all the feature should be considered when training the model. Given this, a model with the intent of predicting the displacement of the points in the z coordinate axis was trained. Its results are compared to the prediction result of RF and represented in Figure 6.5.

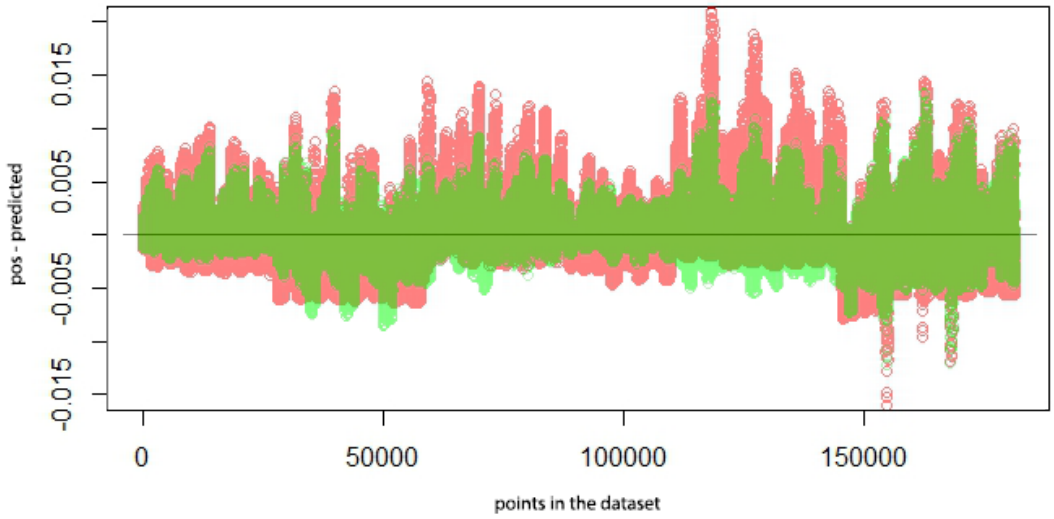


Figure 6.5: Comparison between the predictions of RF and MLP regression models. The y axis represents the distance between the point in the pos-surgical model and the predicted model, in the z axis of the Cartesian coordinate system. The x axis represents all the points of the dataset, considering all the 288 patients. The distance relative the RF model is displayed in green, while the distance of the MLP model is displayed in red.

As represented in Figure 6.5, the distance between the pos-surgical and the predicted models of the breast's patients is significantly larger when using MLP instead of RF.

6.2.3 Multi-output Regressor Results

Given the unsatisfactory results of MLP, the following experiments will be using a more promising algorithm such as Multi-output regression (MOR) algorithms. This type of algorithms usually lead to better results than RF and allow to predict several target variables in the same model, unlike what was done so far that for each axis, a different RF model was used.

Using MOR led to the development of four new scenarios, being all of them tested with both LOO and random train/test splits of the data. With the first and second scenarios, the model tries to predict the displacement of the points in the three axis of the Cartesian coordinate system. In

the first scenario, only the surface points of the breast's model were used to train the model. The evaluation metrics regarding this scenario may be found in Table 6.14 and Table 6.15.

			predicted to pos	pos to predicted	pre to pos	pos to pre	predicted to pre	pre to predicted
3D	**Euclidean Distance**	**Mean**	1.227	1.225	1.758	1.731	1.934	1.950
		Standard Deviation	0.863	0.857	1.333	1.277	1.194	1.226
	Hausdorff Distance		4.336	5.280	6.513	6.317	5.382	5.555
1D	**Euclidean Distance**	**Mean**	0.114	0.114	0.160	0.110	0.127	0.173
		Standard Deviation	0.121	0.128	0.346	0.110	0.121	0.355
	Hausdorff Distance		0.939	1.096	3.168	0.704	0.747	3.403

Table 6.14: Global Evaluation Metrics for MOR model considering only breast surface points to predict the displacement of the points in the three different axis. This results are relative to the LOO train/test split.

			predicted to pos	pre to pos	predicted to pre
3D	**Euclidean Distance**	**Mean**	1.298	2.206	2.224
		Standard Deviation	0.983	1.920	1.523
	Hausdorff Distance		4.857	8.410	6.115
1D	**Euclidean Distance**	**Mean**	1.000	1.904	2.062
		Standard Deviation	0.968	1.929	1.563
	Hausdorff Distance		4.498	8.092	5.999

Table 6.15: Local Evaluation Metrics for MOR model considering only breast surface points to predict the displacement of the points in the three different axis. This results are relative to the LOO train/test split.

The second scenario was trained in the same condition, however using also internal points of the breasts' 3D models. The evaluation metrics are represented in Table 6.16 and Table 6.17.

By comparing the evaluation metrics of these scenarios with the correspondent trials using RF, it is possible to perceive that predicting the 3 variables simultaneously led to worse results. A study short study of the points' behaviour was done and is represented in Figure 6.6. This study allowed to understand how MOR could be improved.

By analysing the displacement of the points in the three different axis of the Cartesian coordinate system, and knowing that the same value of x on all the charts in Figure 6.6, represent the same point of the dataset and consequently the same patient with the same properties, the behaviour of the points on x and y axis seems widely correlated.

Considering the new findings, two more scenarios similar to the previous ones, were created using MOR. Unlike in the previous scenarios, two models will be used: one using MOR for

			predicted to pos	pos to predicted	pre to pos	pos to pre	predicted to pre	pre to predicted
3D	**Euclidean Distance**	**Mean**	1.143	1.142	1.555	1.529	1.779	1.798
		Standard Deviation	0.874	0.870	1.382	1.319	1.286	1.327
	Hausdorff Distance		4.997	4.924	7.107	6.406	5.706	4.924
1D	**Euclidean Distance**	**Mean**	0.060	0.061	0.080	0.055	0.064	0.092
		Standard Deviation	0.067	0.081	0.236	0.060	0.067	0.274
	Hausdorff Distance		0.626	1.060	3.155	0.477	0.494	3.751

Table 6.16: Global Evaluation Metrics for MOR model considering surface and internal points of the breast's 3D model to predict the displacement of the points in the three different axis. This results are relative to the LOO train/test split.

			predicted to pos	pre to pos	predicted to pre
3D	**Euclidean Distance**	**Mean**	1.182	1.919	2.029
		Standard Deviation	0.953	1.940	1.624
	Hausdorff Distance		5.492	9.510	6.694
1D	**Euclidean Distance**	**Mean**	0.933	1.693	1.905
		Standard Deviation	0.930	1.908	1.629
	Hausdorff Distance		5.134	9.147	6.544

Table 6.17: Local Evaluation Metrics for MOR model considering surface and internal points of the breast's 3D model to predict the displacement of the points in the three different axis. This results are relative to the LOO train/test split.

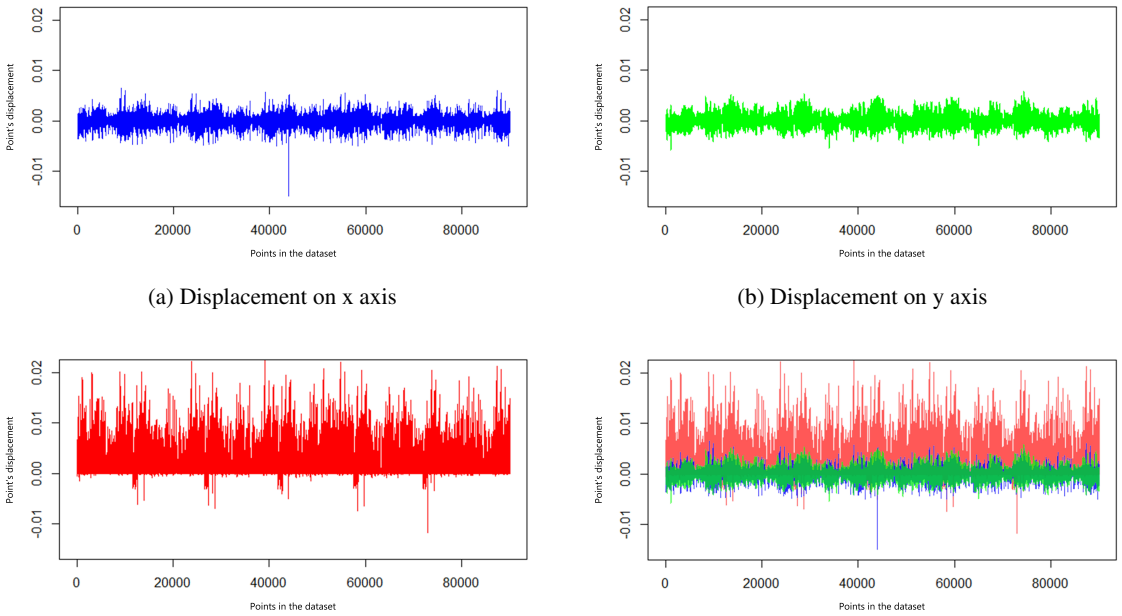
predicting the displacement in x and y axis; and one model using RF to predict the displacement in z axis.

Table 6.18 and Table 6.19 describe the evaluation metrics results for the last described attempt using as input the information from the surface points of the breast's models. The same scenario was also performed considering both the information of the surface points of the breast models and the internal points of the same 3D models. These results are described in Table 6.20 and Table 6.21.

Despite of the good results presented on Table 6.18 and Table 6.20, the evaluation metrics regarding the models trained using RF are slightly better.

6.3 Predictive errors heatmap

In spite of the good results regarding the mean euclidean distance between the points, the still significant value of the Hausdorff distance points out that some of the points are not moving as much as they should. Analysing the points of the breast were this generally happens would make



(a) Displacement on x axis

(b) Displacement on y axis

(c) Displacement on z axis

(d) Comparison between the displacements on the three axis

Figure 6.6: Displacement between the pos and pre-surgical 3D models for all the points of the patients in the dataset. The y values represent the displacement in meters of each point of each patient, represented in x. Being equally order we can assume that the same value on x in any image represents the same point on the dataset.

			predicted to pos	**pos to predicted**	**pre to pos**	**pos to pre**	**predicted to pre**	**pre to predicted**
3D	**Euclidean Distance**	**Mean**	1.175	1.173	1.758	1.731	1.883	1.897
		Standard Deviation	0.851	0.844	1.333	1.277	1.181	1.208
	Hausdorff Distance		4.277	4.214	6.513	6.317	5.246	5.418
1D	**Euclidean Distance**	**Mean**	0.106	0.105	0.160	0.110	0.120	0.162
		Standard Deviation	0.117	0.118	0.346	0.110	0.116	0.330
	Hausdorff Distance		0.953	0.964	3.168	0.704	0.711	3.166

Table 6.18: Global Evaluation Metrics considering surface points of the breast 3D model to predict the displacement of the points in x and y axis (using MOR) and the displacement in z axis (using RF). This results are relative to the LOO train/test split.

us understand what parts of the breast's shape are more unreliable. This predictive errors are calculated based on the distance between the predicted point and the correspondent point on the pos-surgical model of the breast. As they can be seen in Figure 6.7, the predictive errors are displayed as a heatmap representation in a shape of a breast. The representation of the point in a colour near to red indicates a more inaccurate prediction of the point's displacement.

			predicted to pos	pre to pos	predicted to pre
3D	**Euclidean Distance**	**Mean**	1.241	2.206	2.186
		Standard Deviation	0.966	1.920	1.518
	Hausdorff Distance		4.767	8.410	5.938
1D	**Euclidean Distance**	**Mean**	0.924	1.904	1.981
		Standard Deviation	0.942	1.929	1.565
	Hausdorff Distance		4.369	8.092	5.830

Table 6.19: Local Evaluation Metrics considering surface points of the breast 3D model to predict the displacement of the points in x and y axis (using MOR) and the displacement in z axis (using RF). This results are relative to the LOO train/test split.

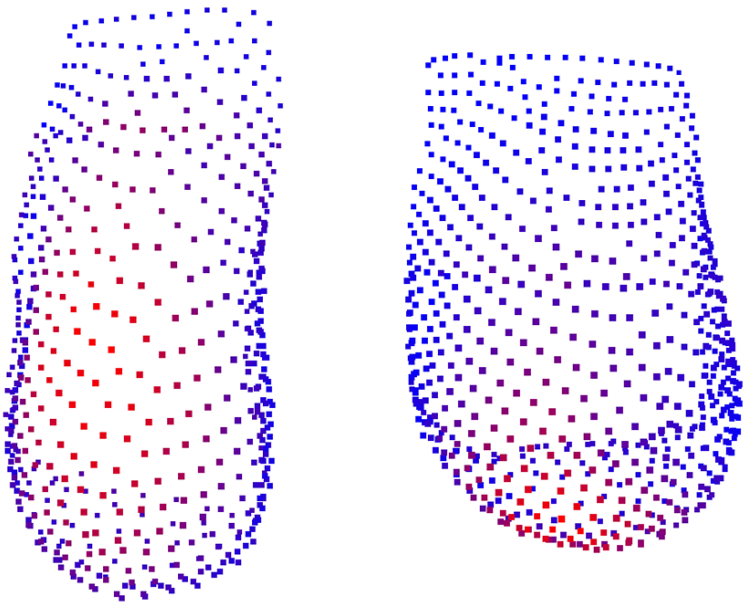
			predicted to pos	pos to predicted	pre to pos	pos to pre	predicted to pre	pre to predicted
3D	**Euclidean Distance**	**Mean**	1.071	1.069	1.555	1.529	1.689	1.705
		Standard Deviation	0.846	0.841	1.382	1.319	1.251	1.287
	Hausdorff Distance		4.941	4.846	7.107	6.406	5.465	5.835
1D	**Euclidean Distance**	**Mean**	0.055	0.056	0.080	0.055	0.060	0.083
		Standard Deviation	0.063	0.068	0.236	0.060	0.062	0.239
	Hausdorff Distance		0.646	0.794	3.155	0.477	0.468	3.330

Table 6.20: Global Evaluation Metrics considering surface and internal points of the breast 3D model to predict the displacement of the points in x and y axis (using MOR) and the displacement in z axis (using RF). This results are relative to the LOO train/test split.

			predicted to pos	pre to pos	predicted to pre
3D	**Euclidean Distance**	**Mean**	1.105	1.919	1.936
		Standard Deviation	0.920	1.940	1.584
	Hausdorff Distance		5.439	9.510	6.441
1D	**Euclidean Distance**	**Mean**	0.828	1.693	1.763
		Standard Deviation	0.898	1.908	1.598
	Hausdorff Distance		5.076	9.147	6.284

Table 6.21: Local Evaluation Metrics considering surface and internal points of the breast 3D model to predict the displacement of the points in x and y axis (using MOR) and the displacement in z axis (using RF). This results are relative to the LOO train/test split.

As expected, the parts of the breast where a greater number of predictive errors exist are coincident with the regions of the breast near to the tumor's location. This regions are more likely



(a) (b)

Figure 6.7: Prediction errors heatmap

to present larger errors, since they are the most affected ones by the deformations caused by the BCS. The points of the breast that are closer to what would be the pectoral muscle, do not present errors.

6.4 Summary

This chapter presented the result of the evaluation metrics defined in chapter 5 for the different methods though to be able to predict the deformations of the breast caused by BCS. From the analysis of the same results, conclusions were able to be drawn and are stated during this chapter.

Regarding the overall results, it is possible to conclude that machine learning techniques led to significantly better results than the methods based on common sense and the feature analysis findings. Despite of the prediction error and the need to minimize them, the results show that the prediction of the breast shape on the proposed environment can be done using machine learning techniques and replace the need to use highly time and computational power consuming alternatives like FEM.

Chapter 7

Conclusions

Due to the high survivability of breast cancer, treatments that lead to a better QoL and a better aesthetic outcomes are being more popular. In order to offer these conditions to the patients, health professional can take advantage from tools in order to assists the patients to choose and follow the most appropriated treatment. The present dissertation focused on the application of machine learning techniques in order to allow to use the deformation predictions caused by BCS in a real-time scenario. Alongside with the dataset generation, a planning tool was developed, that would allow a Health professional to position the tumor and define its properties on a 3D representation of the patient's own 3D model of the breast. Such application was also useful for producing the dataset. Consequently a study on the clinical features was done in order to understand the behaviour of the breast on the wound healing process. At the final stage of the dissertation, following the definition of the models' evaluation metrics, some machine learning approaches such as RF, MLP and MOR were addressed.

The developed work allowed to predict breast deformations caused by BCS in only a small portion of the time that would take a FEM approach to predict the same deformation. Based on the achieved results, and despite of the necessary improvements, machine learning techniques lead to significantly good results and when slightly improved will be able to completely replace the time consuming and highly computational power requirements alternatives, such as FEM.

7.0.1 Future Work

Regarding the planning tool, and despite of the input of some health professionals on what application features the tool should include, the developed one lacks of acceptance and usability tests targeted to the medical community. Through the results and conclusion drawn from such tests, some interface adjustment may need to be made.

The dataset generation was a very exhaustive process and to the lack of more clinical MRI data to generate 3D models, the creation of a larger dataset was not possible. Having more data

would be useful for a better training of the machine learning models and would allow to try some deep learning algorithms.

Concerning the used ML approaches, there are still some possibilities that should be tested. The designed models used as entries individual points and some additional features. An interesting alternative would be using the whole point cloud as an entry of the model. The results presented in chapter 6, despite of the small errors on the mean distance between the predicted and the pos-surgical models, still present an unsatisfactory error on the maximum distance between the models. This could be reduced by changing the objective function of the models' cross validation in order to minimize to minimize the maximum prediction error, instead of fitting the model according to the root mean squared errors.

Also, it would also be interesting to predict cumulative deformations for shorter periods of time.

References

[AAAAAR+14] Mohammed Al-Azri, Huda Al-Awisi, Samira Al-Rasbi, Kawther El-Shafie, Mustafa Al-Hinai, Hamdan Al-Habsi, and Mansour Al-Moundhri. Psychosocial impact of breast cancer diagnosis among Omani women. *Oman Medical Journal*, 29(6):437–444, 2014.

[AMS76] F S Azar, D Metaxas, and M D Schnall. A Finite Element Model of the Breast for Predicting Mechanical Deformations during Interventional Procedures. 00(c):5, 1976.

[BHR+17] Thore M. Bücking, Emma R. Hill, James L. Robertson, Efthymios Maneas, Andrew A. Plumb, and Daniil I. Nikitichev. From medical imaging data to 3d printed anatomical models. *PLOS ONE*, 12(5):1–10, 05 2017.

[BI90] R Berwick and Village Idiot. An Idiot ' s guide to Support vector machines (SVMs) SVMs : A New Generation of Learning Algorithms Key Ideas. pages 1–28, 1990.

[Bio] Computational Biomechanics. *No Title*.

[BVBM] Hanen Borchani, Gherardo Varando, Concha Bielza, and Boadilla Monte. A survey on multi-output regression. pages 1–27.

[CHO09] CRUI CHOU. Learning Based Coarse-to-fine Image Registration. *Icfcc, 2010*, 2009.

[CMZO14] Pedro Costa, João P. Monteiro, Hooshiar Zolfagharnasab, and Hélder P. Oliveira. Tessellation-based coarse registration method for 3D reconstruction of the female torso. *Proceedings - 2014 IEEE International Conference on Bioinformatics and Biomedicine, IEEE BIBM 2014*, pages 301–306, 2014.

[Cut13] Adele Cutler. Trees and Random Forests. pages 1–92, 2013.

[DD16] Aculdade De and Niversidade Do. Framework for Planing the Aesthetic Result after Breast Surgery. 2016.

[D'O08] Carl J. D'Orsi. Breast Imaging. *Radiographics*, 42(5):xi–xii, 2008.

[EM13] Harold Ellis and Vishy Mahadevan. Anatomy and physiology of the breast. *Surgery (United Kingdom)*, 31(1):11–14, 2013.

[EPJ+08] Medical Engineering, Amaya Perez, Javier Jover, Herrero Alma, and I T Systems. A finite element model to accurately predict real deformations of the breast. (April), 2008.

[EVH+16] Björn Eiben, Vasileios Vavourakis, John H. Hipwell, Sven Kabus, Thomas Buelow, Cristian Lorenz, Thomy Mertzanidou, Sara Reis, Norman R. Williams, Mohammed Keshtgar, and David J. Hawkes. Symmetric Biomechanically Guided Prone-to-Supine Breast Image Registration. *Annals of Biomedical Engineering*, 44(1):154–173, 2016.

[FFA+00] Patrik Finne, Ralf Finne, Anssi Auvinen, Harri Juusela, Jussi Aro, Liisa Määttänen, Matti Hakama, Sakari Rannikko, Teuvo L J Tammela, and Ulf Håkan Stenman. Predicting the outcome of prostate biopsy in screen-positive men by a multilayer perceptron network. *Urology*, 56(3):418–422, 2000.

[FGH+14] Eva Foersterling, Michael Golatta, Andre Hennigs, Sophie Schulz, Geraldine Rauch, Sarah Schott, Christoph Domschke, Florian Schuetz, Christof Sohn, and Joerg Heil. Predictors of early poor aesthetic outcome after breast-conserving surgery in patients with breast cancer: initial results of a prospective cohort study at a single institution. *Journal of surgical oncology*, 110(7):801–806, 2014.

[FSLH13] E. E. Fowler, T. A. Sellers, B. Lu, and J. J. Heine. Breast Imaging Reporting and Data System (BI-RADS) breast composition descriptors: Automated measurement development for full field digital mammography. *Medical Physics*, 40(11):113502, 2013.

[GM97] Sarah F. F. Gibson and Brian Mirtich. A survey of deformable modeling in computer graphics. Technical report, Mitsubishi Electric Research Laboratories, 1997.

[GSTB13] M. Garbey, R. Salmon, D. Thanoon, and B. L. Bass. Multiscale modeling and distributed computing to predict cosmesis outcome after a lumpectomy. *Journal of Computational Physics*, 244:321–335, 2013.

[Gun98] Steve R Gunn. Support Vector Machines for Classification and Regression by. (May):66, 1998.

[IAI+16] R N E Iben, V Asileios V Avourakis, J O H N H H Ipwell, S V E N K Abus, T Homas B Uelow, C Ristian L Orenz, T Homy M Ertzanidou, S A R A R Eis, N Orman R W Illiams, and M Ohammed K Eshtgar. Symmetric Biomechanically Guided Prone-to-Supine Breast Image Registration. 44(1):154–173, 2016.

[JJM13] Ilyang Joo, Kanghun Jeong, and Hyeonjoon Moon. Breast image registration for PET-CT and MR based on 3D surface matching. *International Journal of Bio-Science and Bio-Technology*, 5(6):201–206, 2013.

[JVM+12] Maria Joa, Conny Vrieling, Douglas Macmillan, Dick Rainsbury, Joerg Heil, Eric Hau, and Mohammed Keshtgar. Recommendations for the aesthetic evaluation of breast cancer conservative treatment. pages 629–637, 2012.

[KM04] Tsuneya Kurihara and Natsuki Miyata. Modeling Deformable Human Hands from Medical Images. *Eurographics/ACM SIGGRAPH Symposium on Computer Animation*, pages 355–363, 2004.

[KSR+08] Min Soon Kim, Juliano C. Sbalchiero, Gregory P. Reece, Michael J. Miller, Elisabeth K. Beahm, and Mia K. Markey. Assessment of Breast Aesthetics. *Plastic and Reconstructive Surgery*, 121(4):186e–194e, 2008.

[LRB+13] Angela W C Lee, Vijayaraghavan Rajagopal, Thiranja P. Babarenda Gamage, Anthony J. Doyle, Poul M F Nielsen, and Martyn P. Nash. Breast lesion co-localisation between X-ray and MR images using finite element modelling. *Medical Image Analysis*, 17(8):1256–1264, 2013.

[MWWC] Technische Universität München, Florian Walch, Florian Walch, and Prof Daniel Cremers. Master ' s Thesis in Informatics Deep Learning for Image-Based Localization Master ' s Thesis in Informatics Deep Learning for Image-Based Localization Deep Learning für bildbasierte Lokalisierung.

[NJ16] Patrik Nygren and Michael Jasinski. A Comparative Study of Segmentation and Classification Methods for 3D Point Clouds. 2016.

[OMH83] G. F. Oster, J. D. Murray, and A. K. Harris. Mechanical aspects of mesenchymal morphogenesis. *J Embryol Exp Morphol*, 78(1):83–125, 1983.

[OoBIBP+16] Alberto (Dept. of Mathematics Ortiz, Computer Science / University of Balearic Islands), Francisco Bonnin-Pascual, Emilio Garcia-Fidalgo, and Joan P. Company. Visual Inspection of Vessels by Means of aMicro-Aerial Vehicle: An Artificial Neural Network Approach for Corrosion Detection. *Robot 2015: Second Iberian Robotics Conference*, 418(October):543–555, 2016.

[PCO14] Diogo Pernes, Jaime S. Cardoso, and Hélder P. Oliveira. Fitting of superquadrics for breast modelling by geometric distance minimization. 2014.

[PIS+13] Paolo Patete, Maria Ida Iacono, Maria Francesca Spadea, Giovanna Trecate, Daniele Vergnaghi, Luca Tommaso Mainardi, and Guido Baroni. A multi-tissue mass-spring model for computer assisted breast surgery. *Medical Engineering and Physics*, 35:47–53, 1 2013.

[PMV03] Josien P W Pluim, J. B A Antoine Maintz, and Max A. Viergever. Mutual-information-based registration of medical images: A survey. *IEEE Transactions on Medical Imaging*, 22(8):986–1004, 2003.

[Qi] Charles R Qi. PointNet: Deep Learning on Point Sets for 3D Classification and Segmentation.

[Rea99] D. Rueckert et al. Nonrigid registration using free-form deformations: application to breast MR images. *IEEE Trans. Med. Imag.*, 18(8):712–21, 1999.

[Row00] Role of Breast Reconstructive Surgery in Physical and Emotional Outcomes Among Breast Cancer Survivors gested that conservation or restitution of the breast might mitigate the nega- tive effects of breast cancer on women ' s ectomy , mastectomy alone , or. *J Natl Cancer Inst*, 92(17):1422–1429, 2000.

[SASB08] Vance Y. Sohn, Zachary M. Arthurs, James A. Sebesta, and Tommy A. Brown. Primary tumor location impacts breast cancer survival. *American Journal of Surgery*, 195(5):641–644, 2008.

[SSAG11] P. D. Sreekanth, P. D. Sreedevi, Shakeel Ahmed, and N. Geethanjali. Comparison of FFNN and ANFIS models for estimating groundwater level. *Environmental Earth Sciences*, 62(6):1301–1310, 2011.

[SSNS12] Nafiza Saidin, HAM Sakim, UK Ngah, and IL Shuaib. Segmentation of breast regions in mammogram based on density: a review. *International Journal of Computer Science Issues*, 9(4):108–116, 2012.

[TCL+13] Gary K L Tam, Zhi-quan Cheng, Yu-kun Lai, Frank C Langbein, Yonghuai Liu, David Marshall, Ralph R Martin, Xian-fang Sun, and Paul L Rosin. Registration of 3D Point Clouds and Meshes: A Survey from Rigid to Nonrigid. *IEEE transactionS ON visualization AND computer graphics*, 19(7):1199–1217, 2013.

[Tei] Teixeira, J. - INESC TEC. Multi-view annotation tool for radiology images.

[VEH+16] Vasileios Vavourakis, Bjoern Eiben, John H. Hipwell, Norman R. Williams, Mo Keshtgar, and David J. Hawkes. Multiscale mechano-biological finite element modelling of oncoplastic breast surgery - Numerical study towards surgical planning and cosmetic outcome prediction. *PLoS ONE*, 11(7), 2016.

[VU98] Computer Vision and Image Understanding. A Parametric Deformable Model To Fit Unstructured 3D Data. 1:39–54, 1998.

[Wit04] Lawrence M Witmer. Clinical Anatomy of the. 2004.

[WPA09] Medha V Wyawahare, Pradeep M Patil, and Hemant K Abhyankar. Image Registration Techniques : An overview. *International Journal of Signal Processing, Image Processing and Pattern Recognition*, 2(3):11–28, 2009.

Appendix A

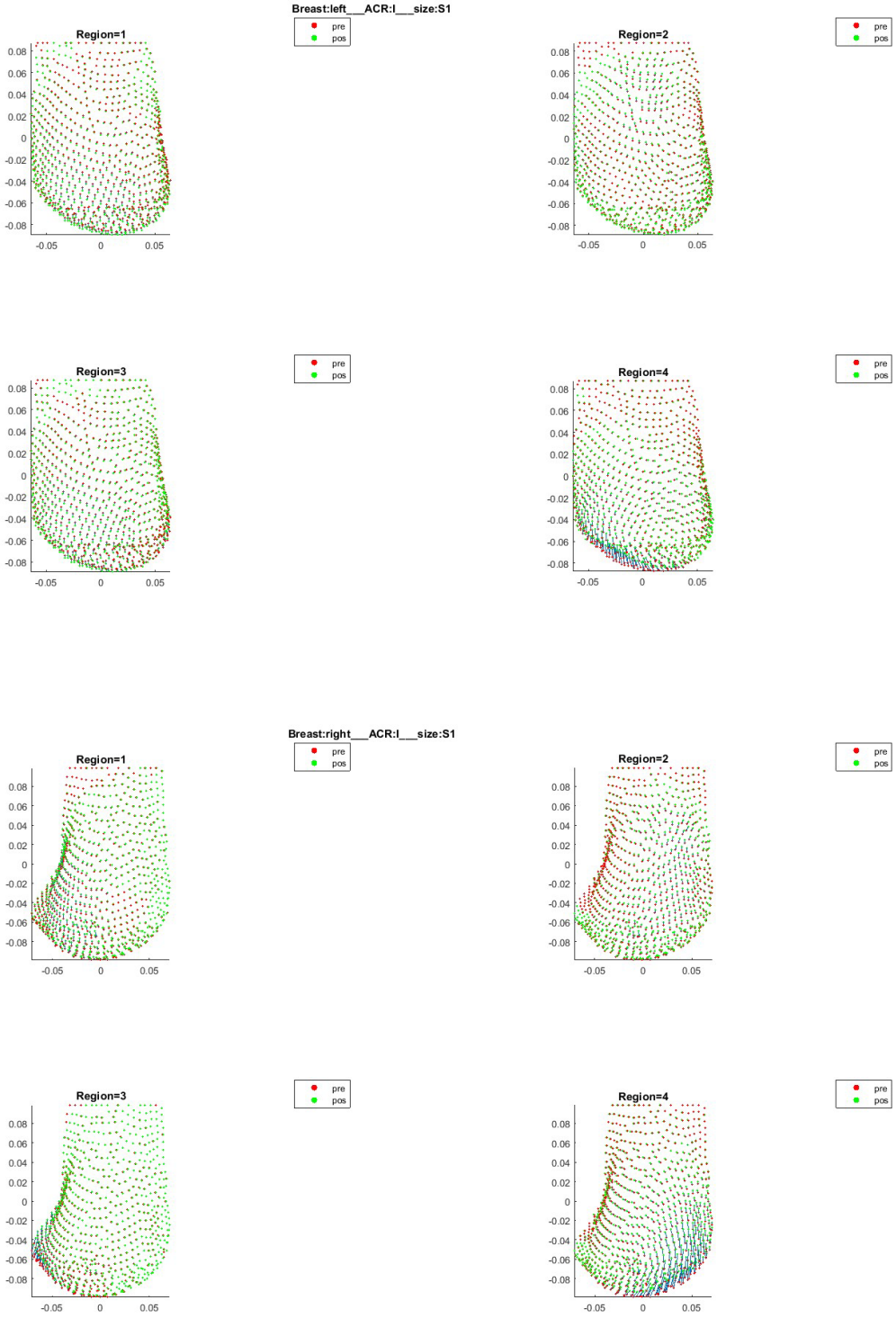
Feature Analysis

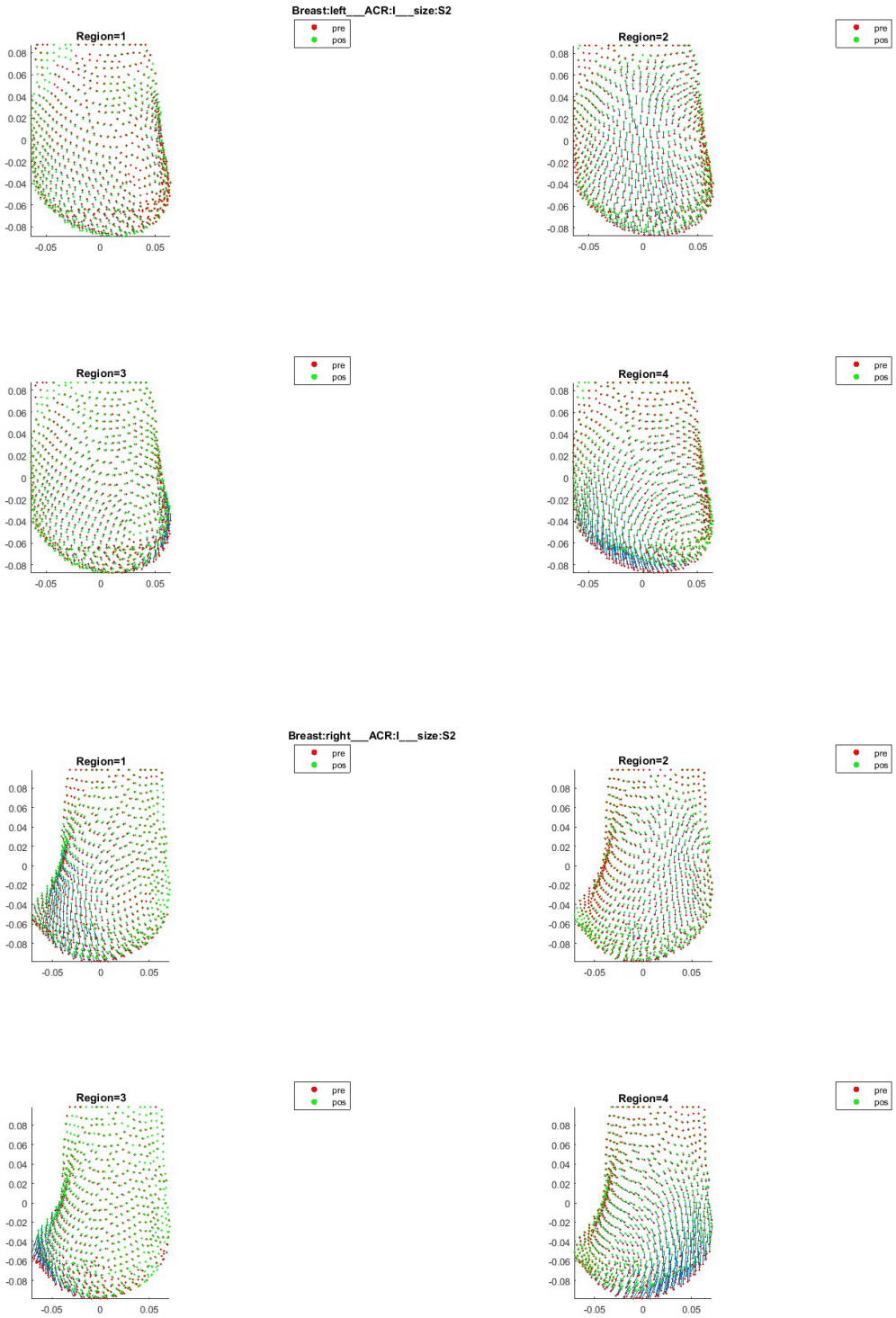
This appendix presents all the results of the feature analysis that was performed in order to understand the impact that the clinical features have on the wound healing simulation.

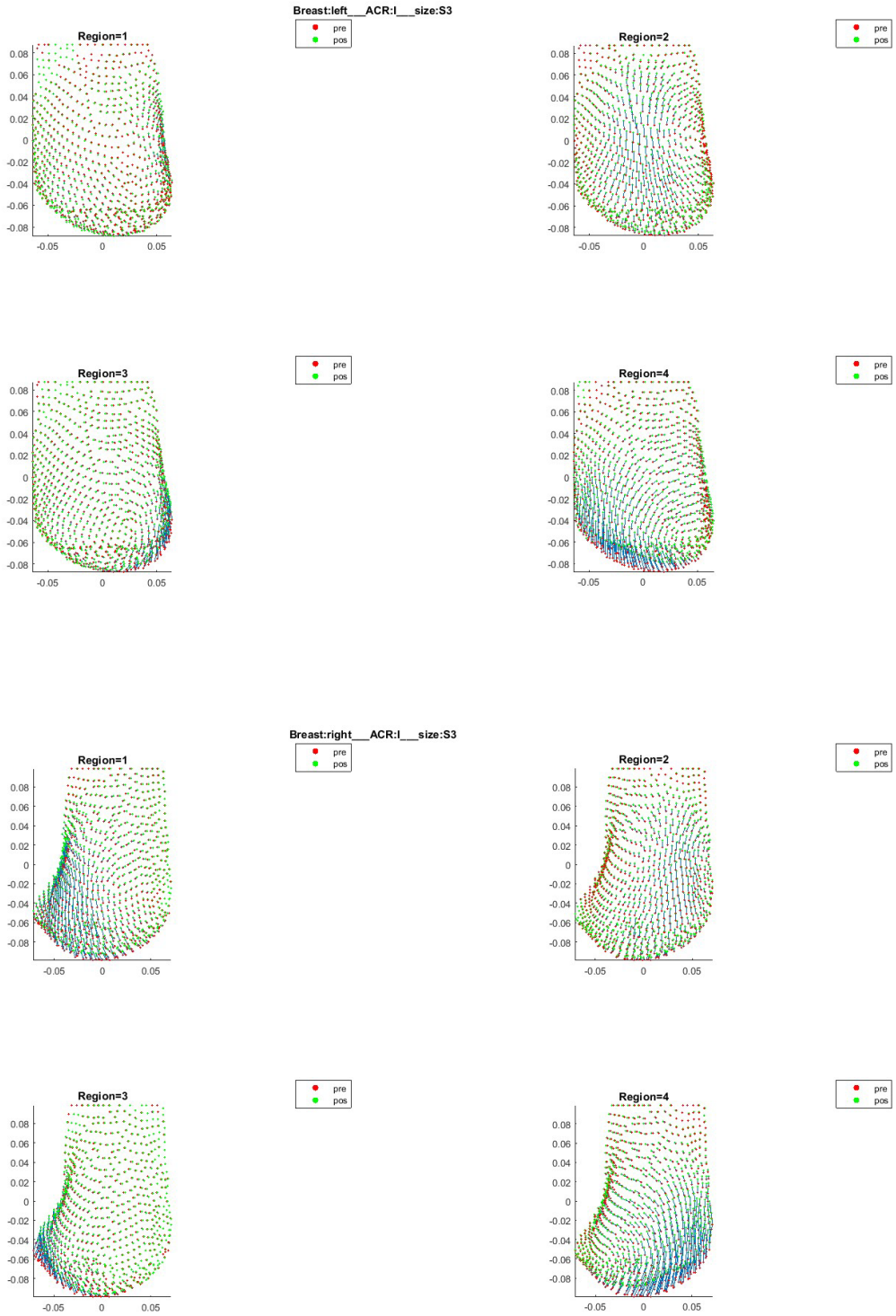
The analysis of the clinical features was done by comparing the displacement of the corresponding points in the pre and pos-surgical models of the breast models between variations of the same patient, where one of the clinical features was changing, and the others were kept the constant. The clinical features that were analysed were the tumor's size, the tumor's region and the density of the breast.

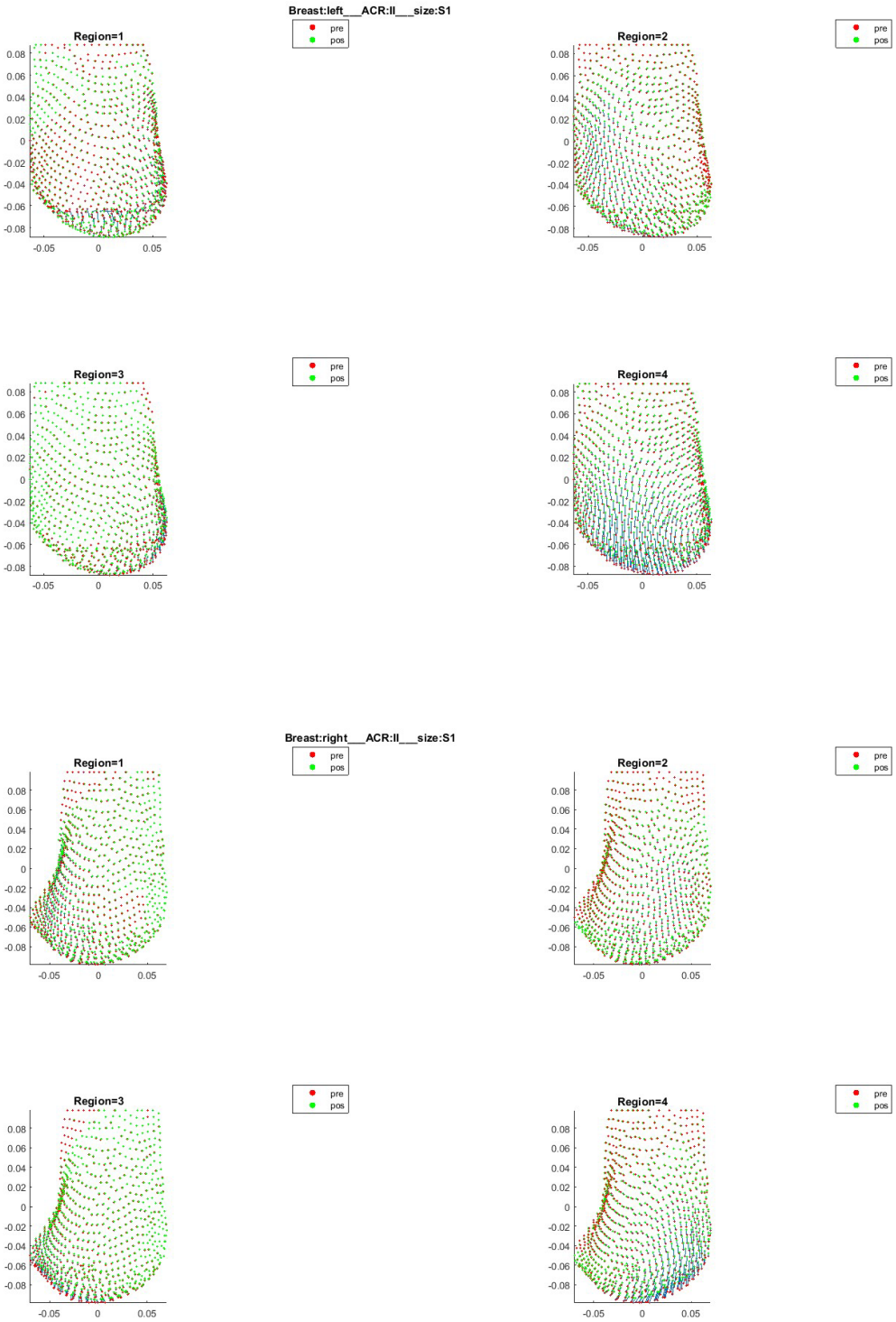
Fixing breast laterality, ACR and tumor size:

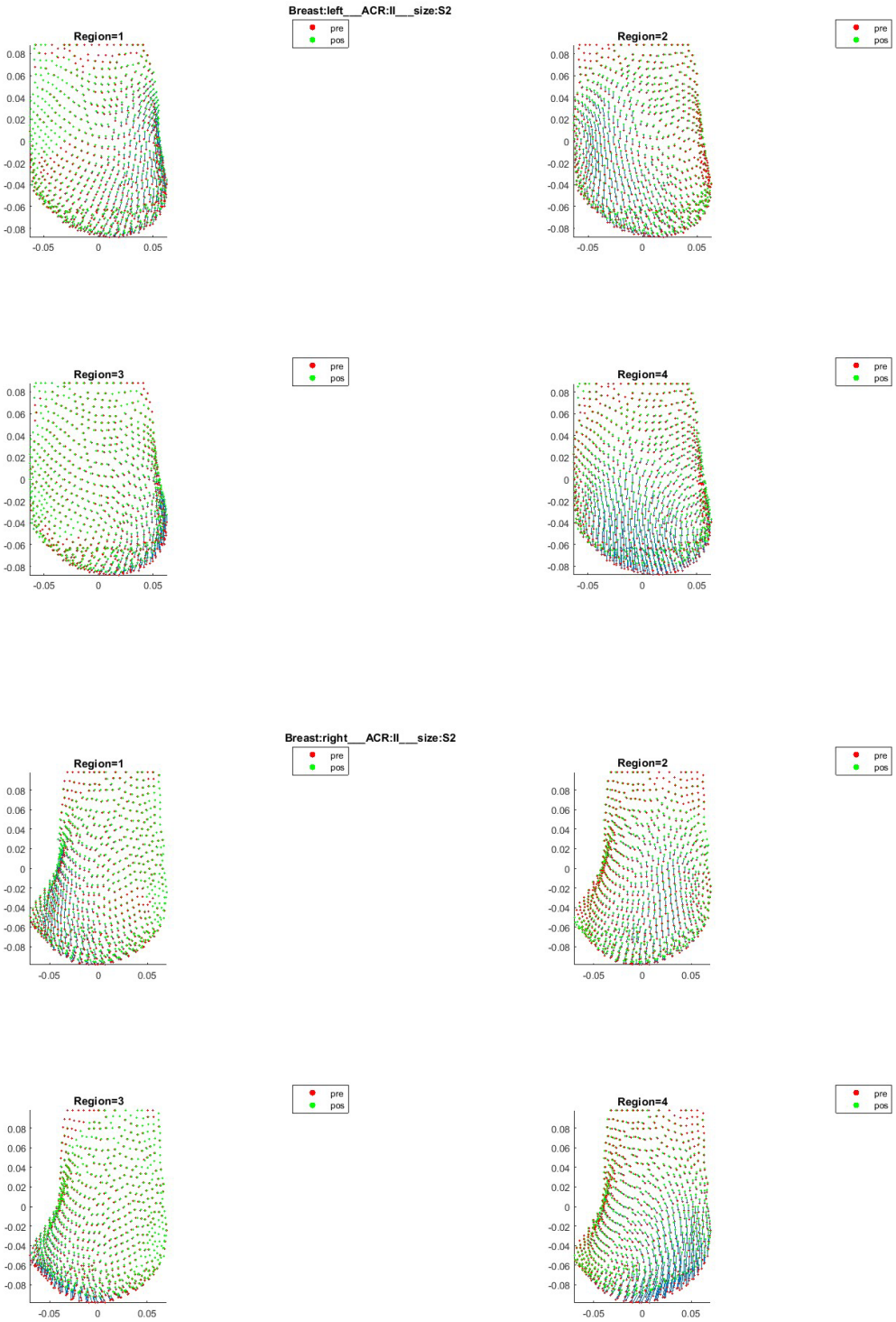
Varying Region

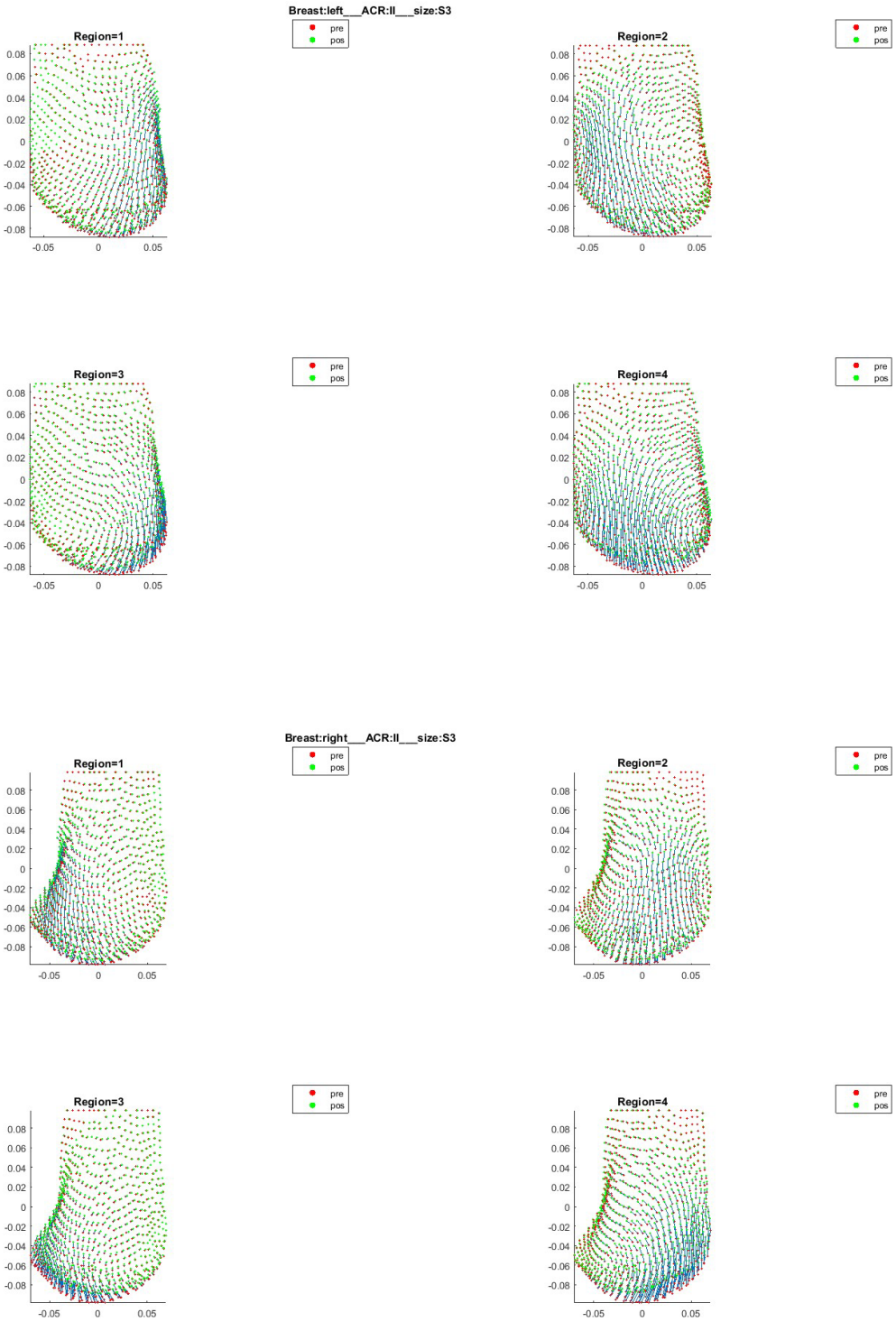


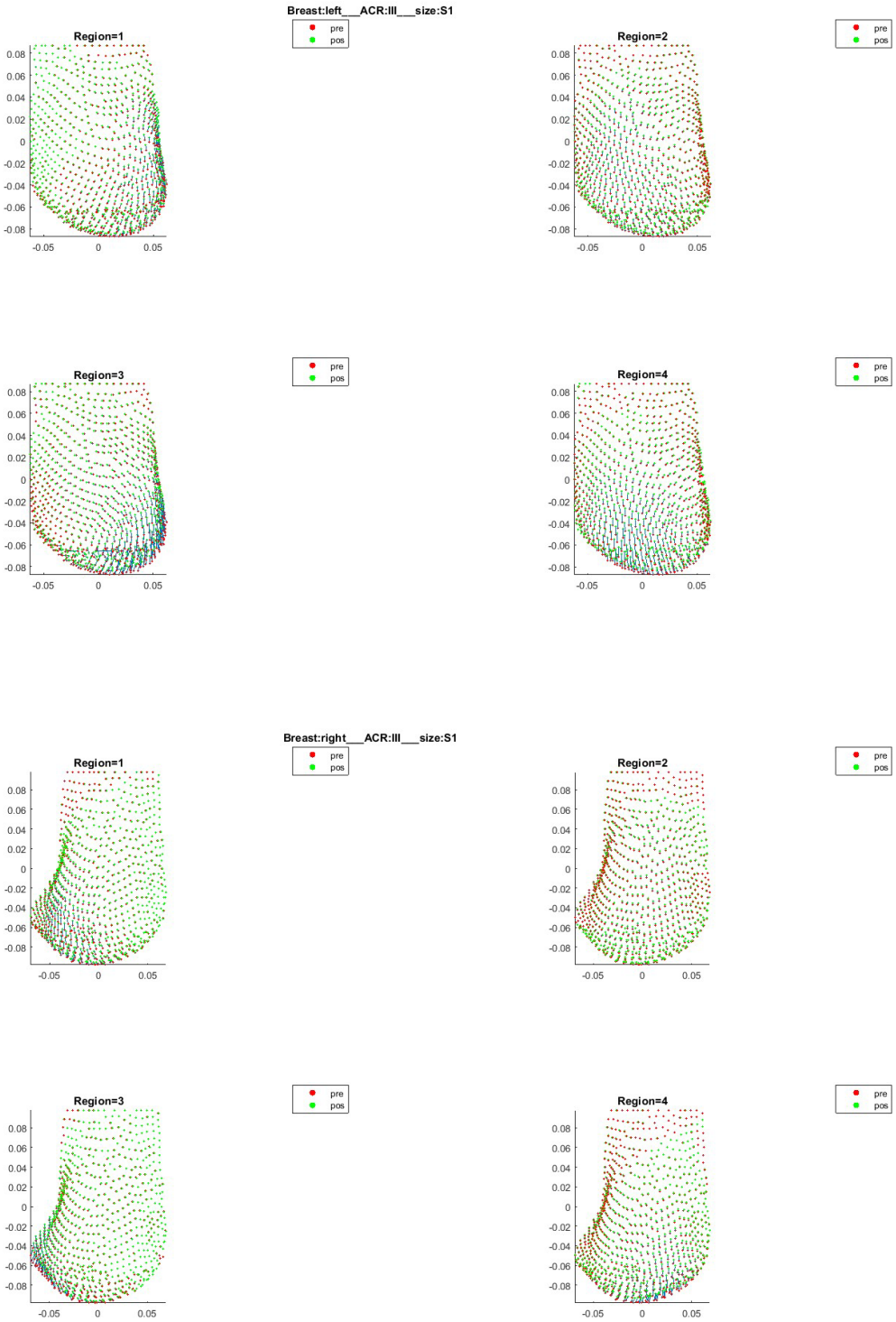

Breast:left___ACR:I___size:S2
Region=1
Region=2
Region=3
Region=4
pre
pos
Breast:right___ACR:I___size:S2
Region=1
Region=2
Region=3
Region=4
pre
pos

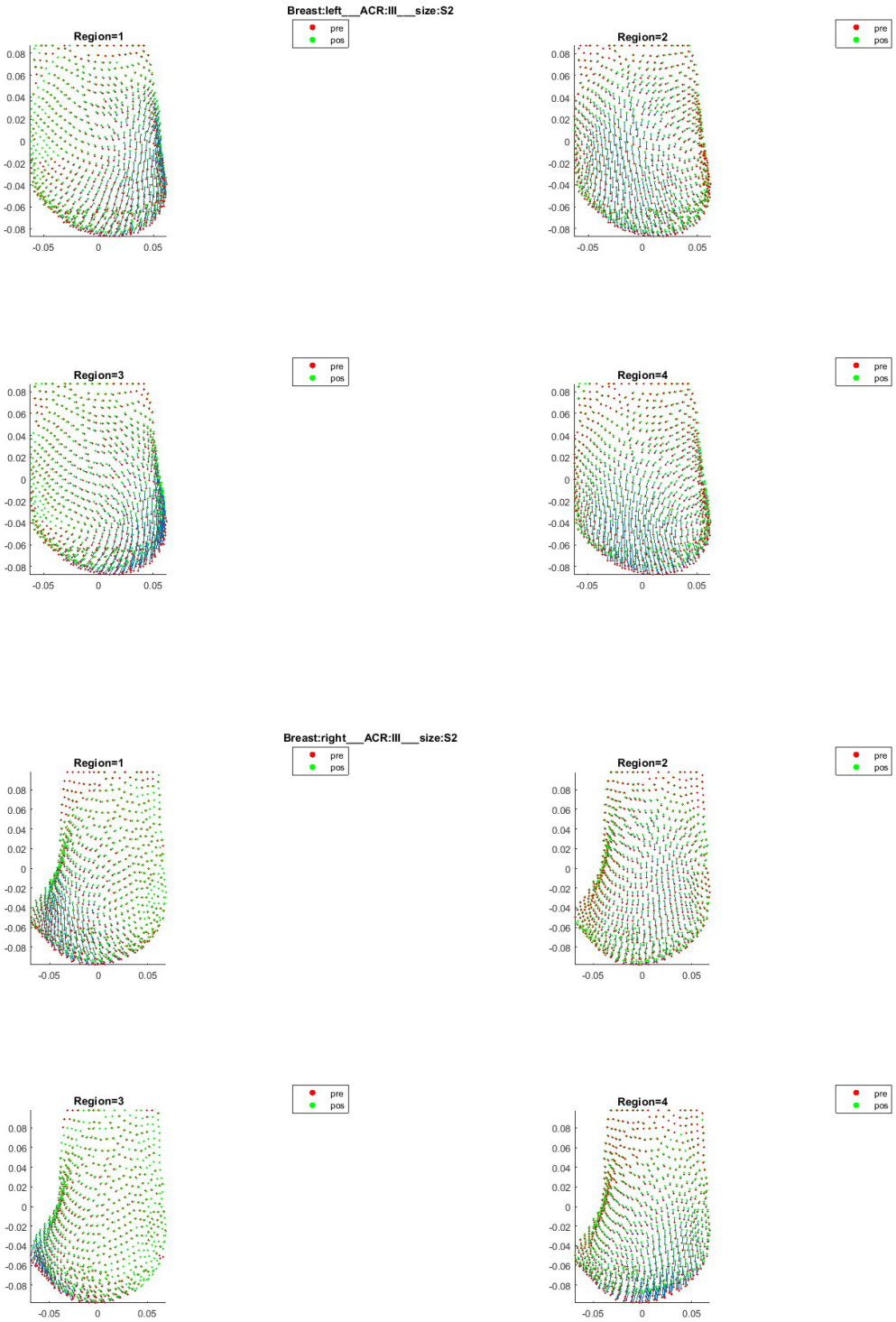

Breast:left___ACR:I___size:S3
Region=1
Region=2
Region=3
Region=4
Breast:right___ACR:I___size:S3
Region=1
Region=2
Region=3
Region=4
pre
pos

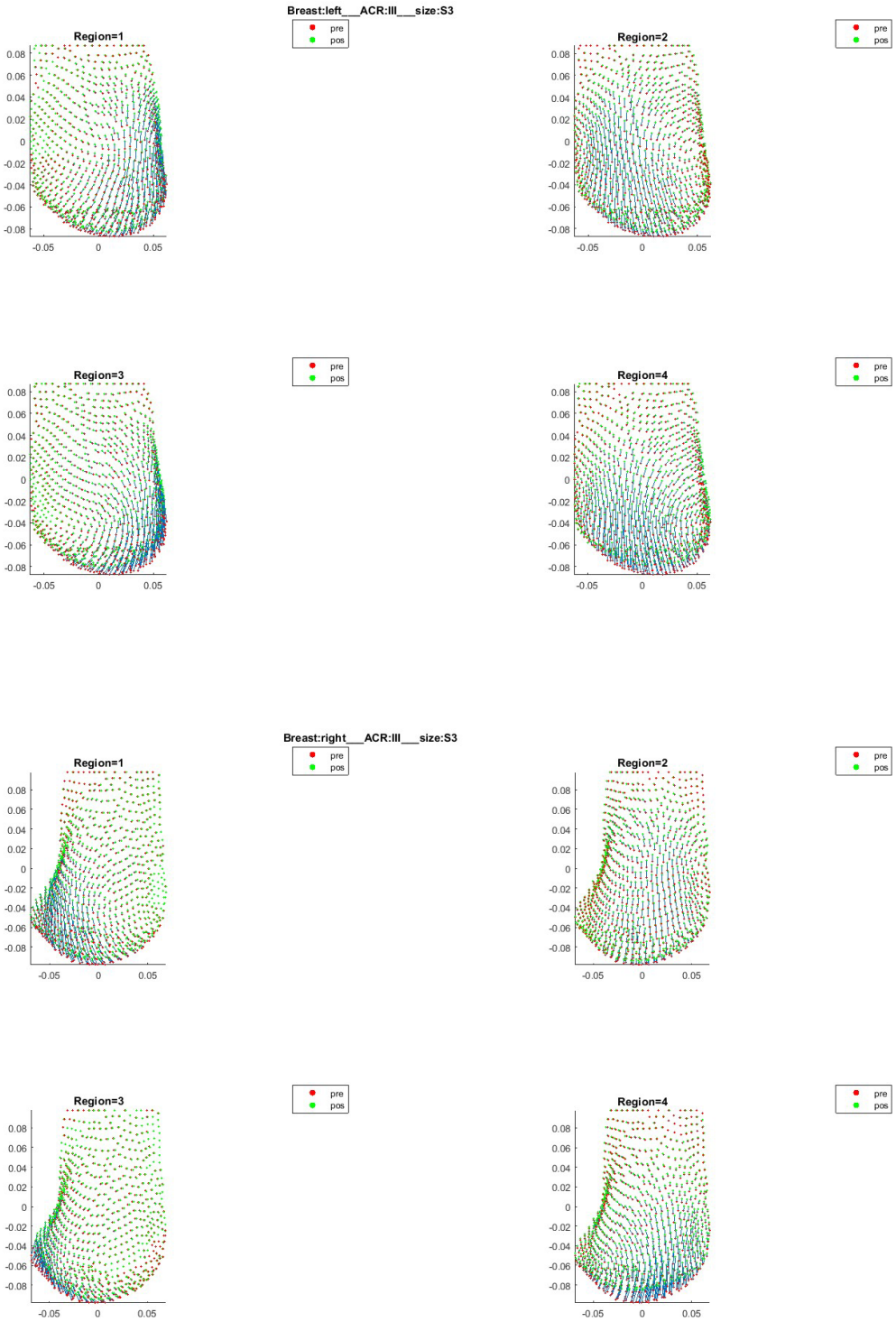

Breast:left___ACR:II___size:S1
Region=1
Region=2
Region=3
Region=4
pre
pos
Breast:right___ACR:II___size:S1
Region=1
Region=2
Region=3
Region=4
pre
pos

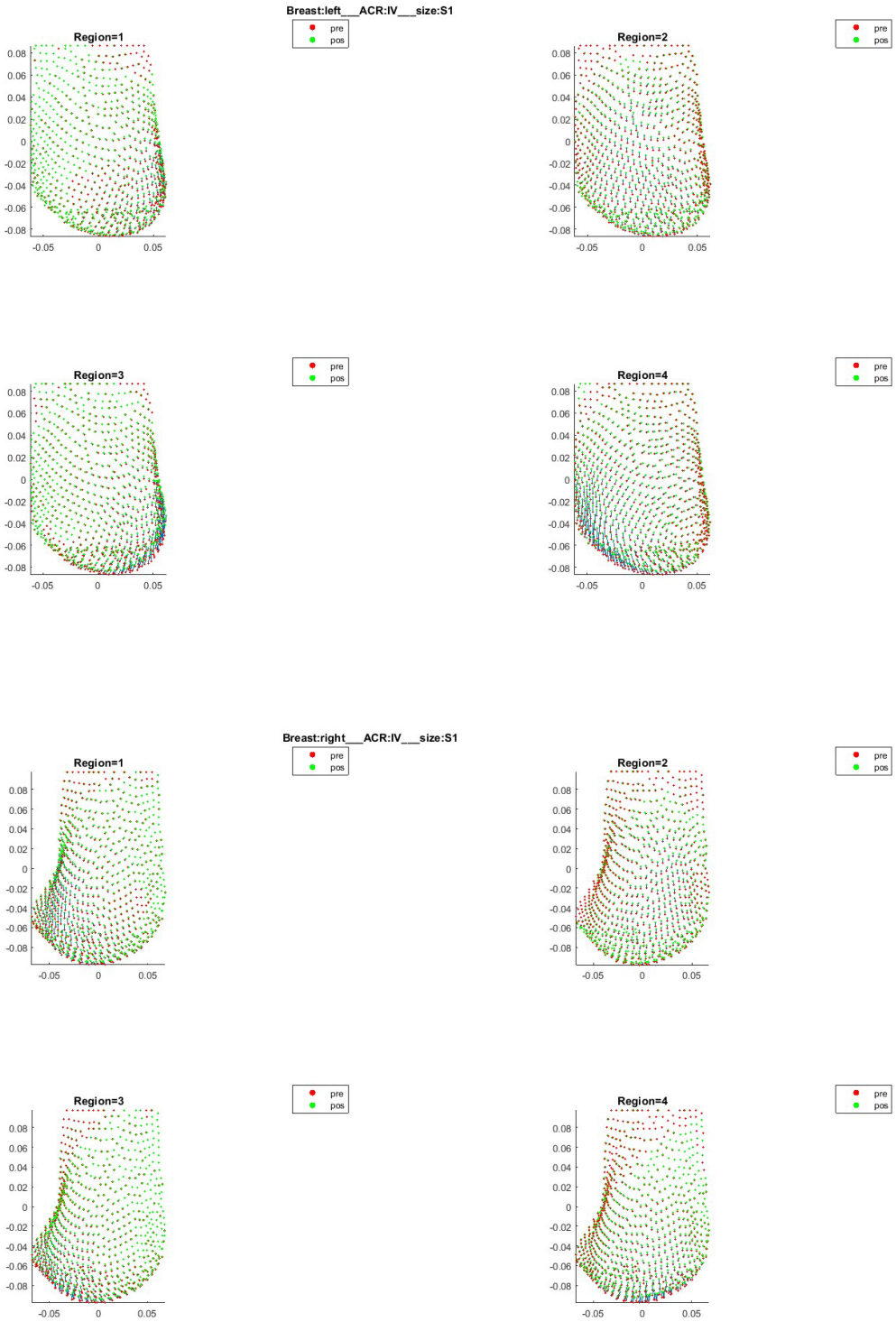

Breast:left___ACR:II___size:S2
Region=1
Region=2
Region=3
Region=4
Breast:right___ACR:II___size:S2
Region=1
Region=2
Region=3
Region=4
pre
pos

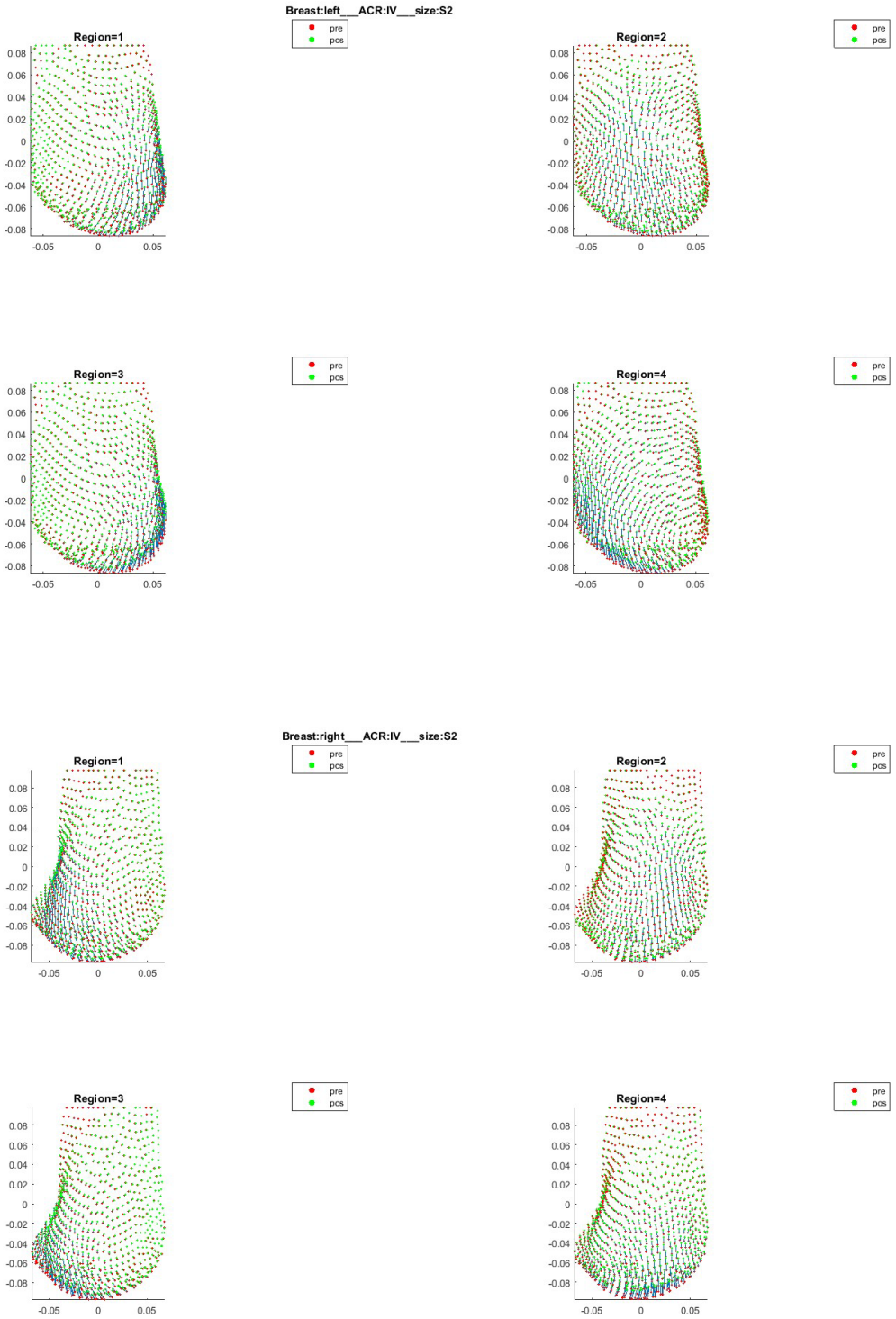

Breast:left___ACR:II___size:S3
Region=1
Region=2
Region=3
Region=4
Breast:right___ACR:II___size:S3
Region=1
Region=2
Region=3
Region=4
pre
pos

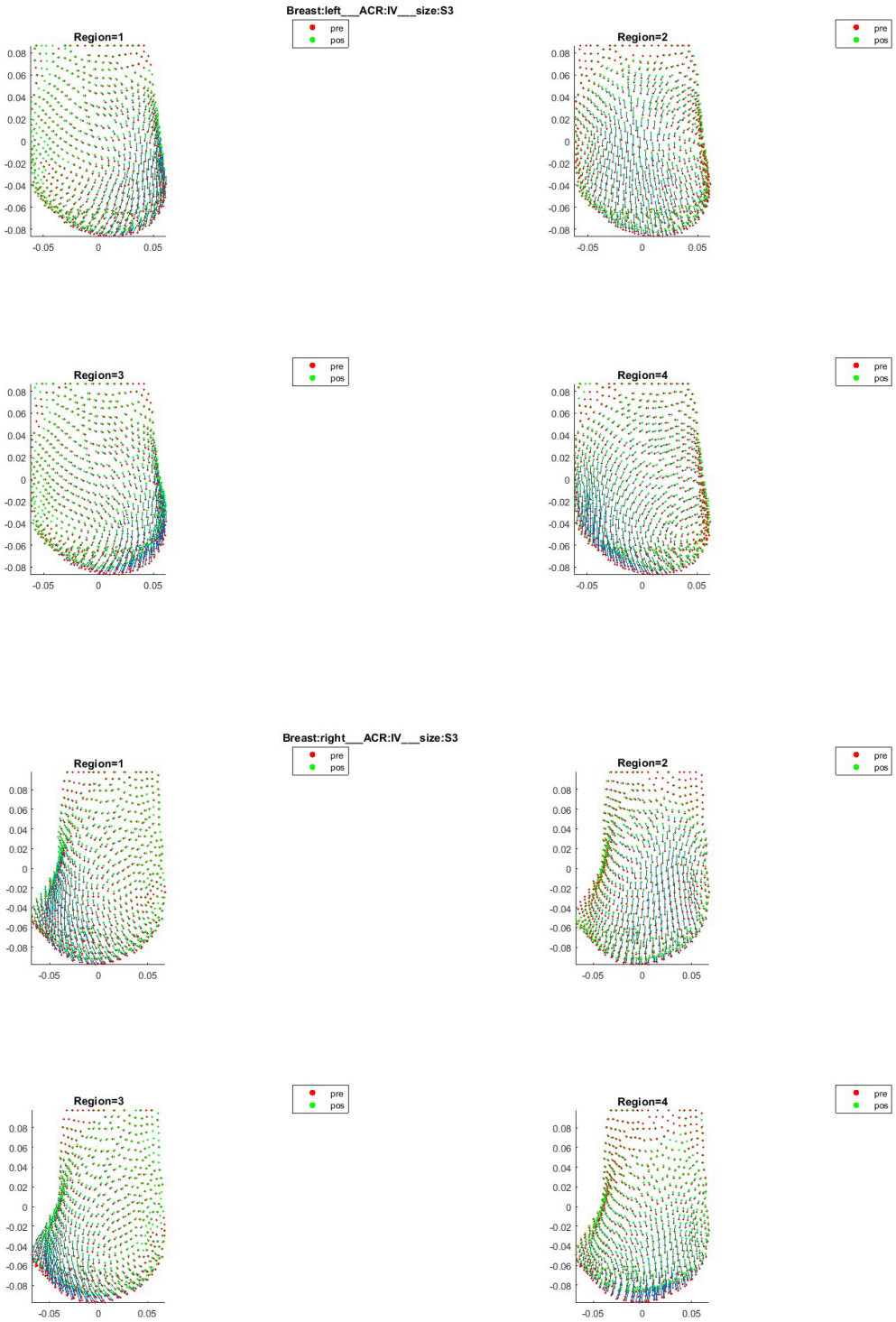

Breast:left___ACR:III___size:S1
Region=1
Region=2
Region=3
Region=4
pre
pos
Breast:right___ACR:III___size:S1
Region=1
Region=2
Region=3
Region=4
pre
pos


Breast:left___ACR:III___size:S2
Region=1
Region=2
Region=3
Region=4
pre
pos
Breast:right___ACR:III___size:S2
Region=1
Region=2
Region=3
Region=4
pre
pos


Breast:left___ACR:III___size:S3
Region=1
Region=2
Region=3
Region=4
pre
pos
Breast:right___ACR:III___size:S3
Region=1
Region=2
Region=3
Region=4
pre
pos

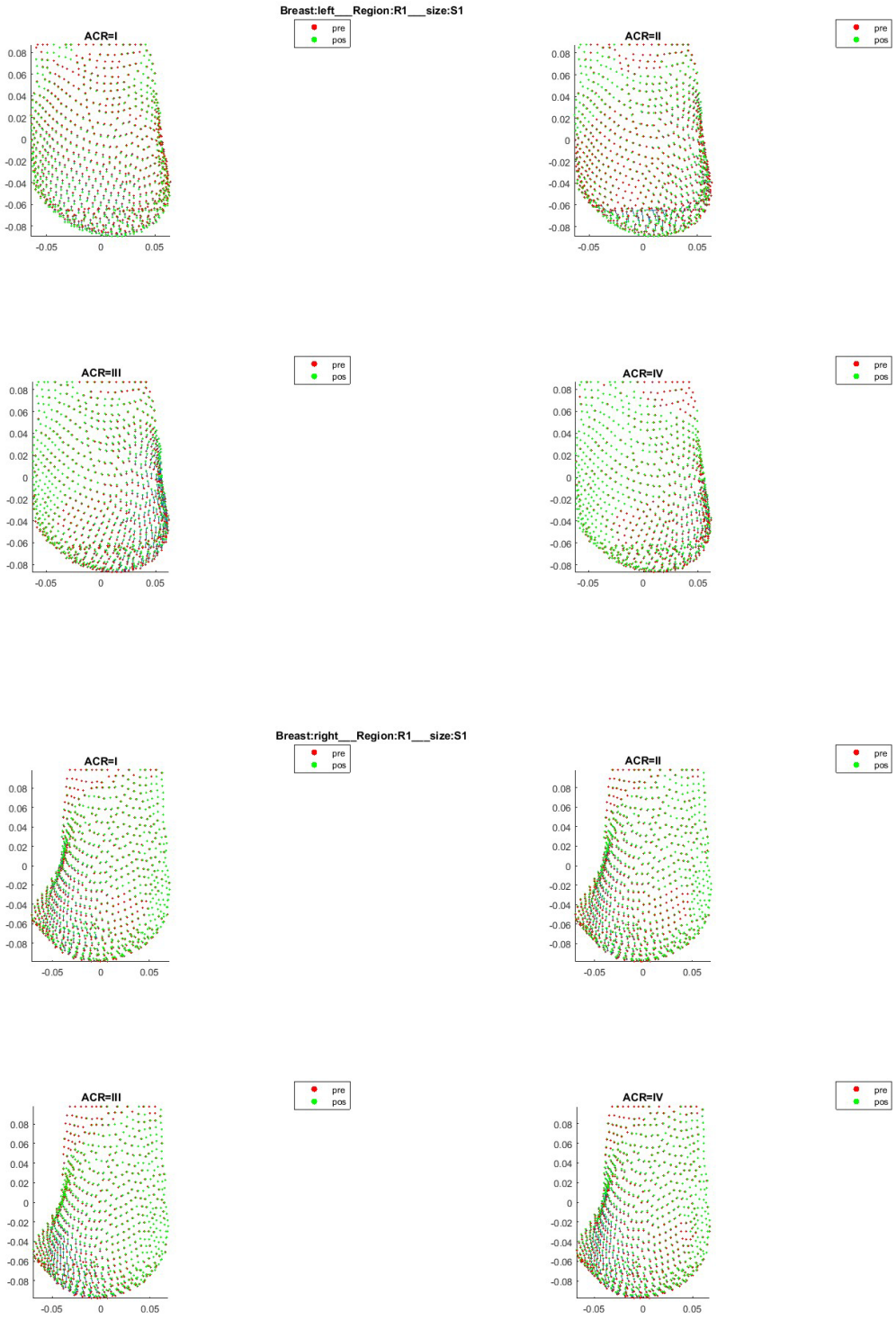

Breast:left___ACR:IV___size:S1
Region=1
Region=2
Region=3
Region=4
pre
pos
Breast:right___ACR:IV___size:S1
Region=1
Region=2
Region=3
Region=4
pre
pos


Breast:left___ACR:IV___size:S2
Region=1
Region=2
Region=3
Region=4
pre
pos
Breast:right___ACR:IV___size:S2
Region=1
Region=2
Region=3
Region=4
pre
pos

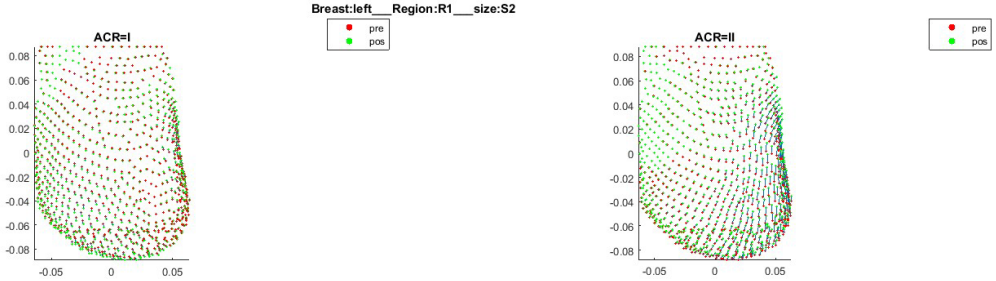

Breast:left___ACR:IV___size:S3
Region=1
Region=2
Region=3
Region=4
pre
pos
Breast:right___ACR:IV___size:S3
Region=1
Region=2
Region=3
Region=4
pre
pos

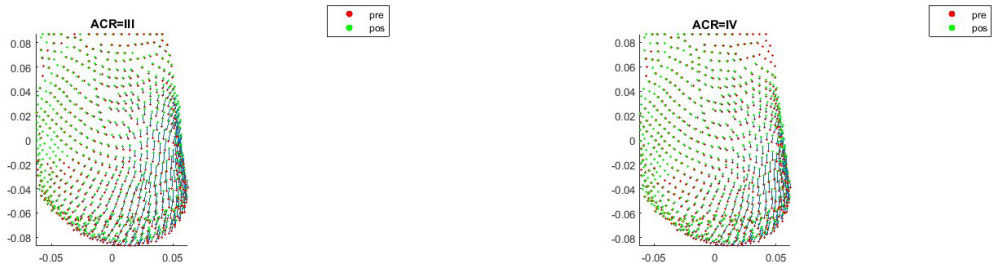
Fixing breast laterality, Region and tumor size:

Varying ACR

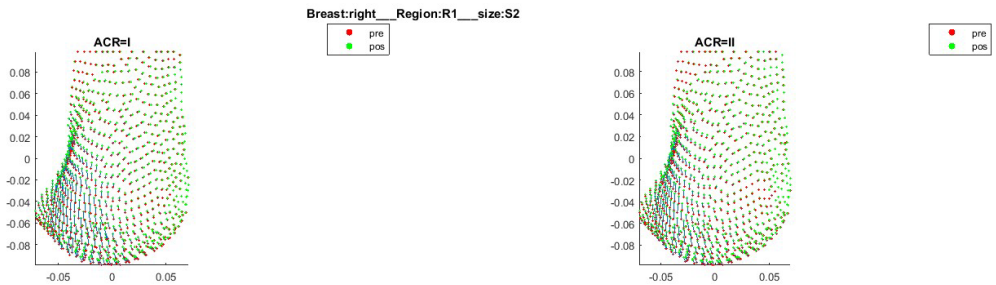


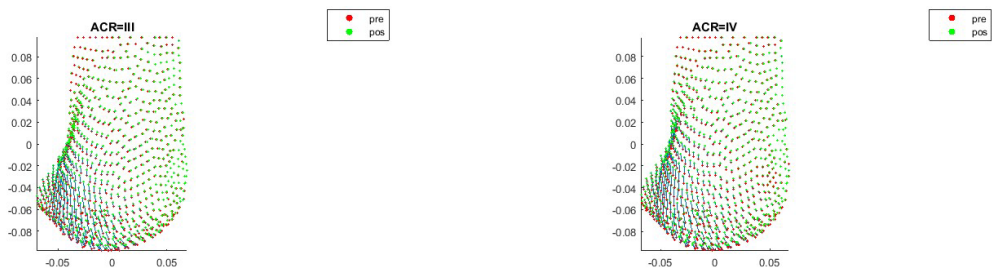
Breast:left___Region:R1___size:S2

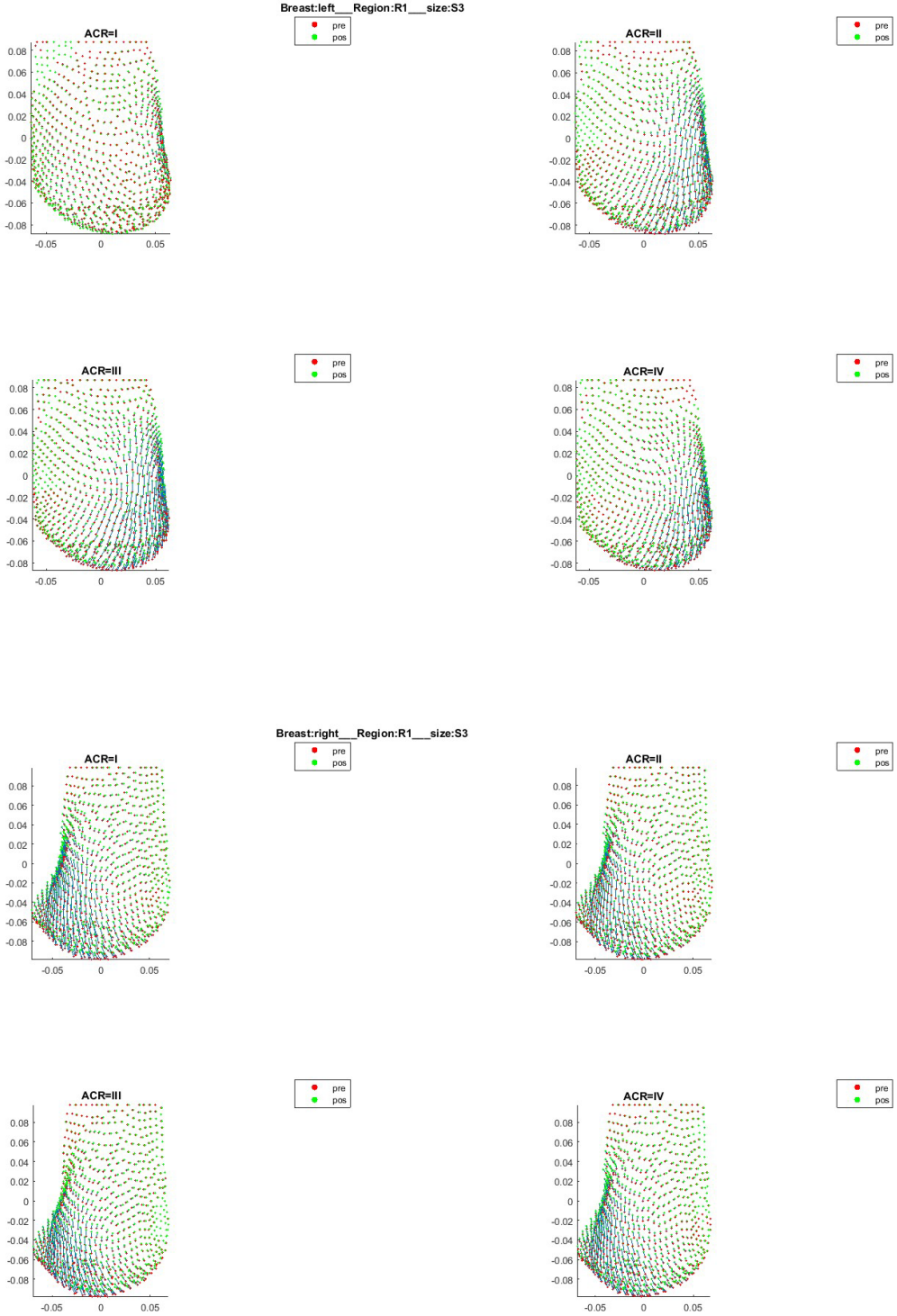


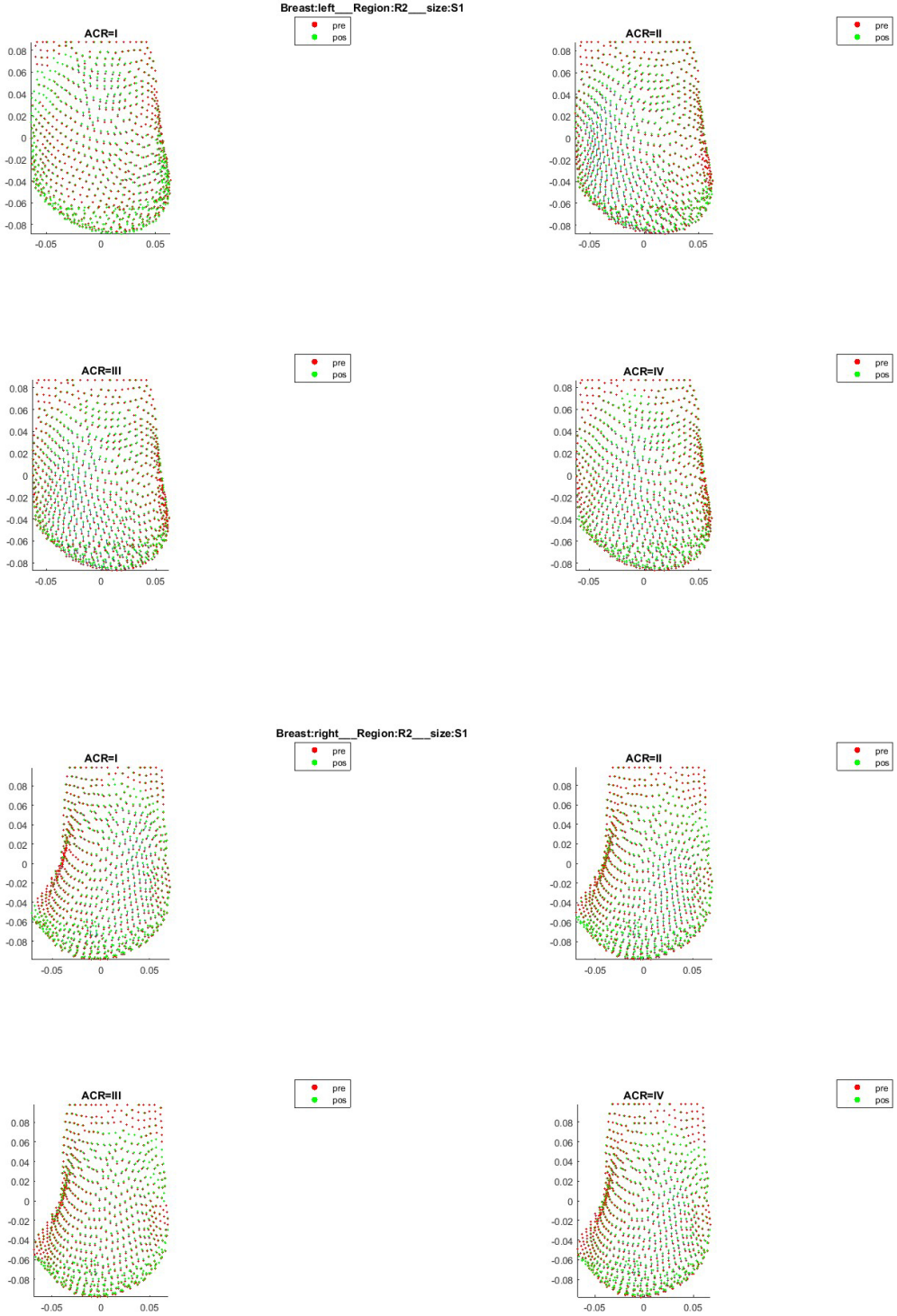


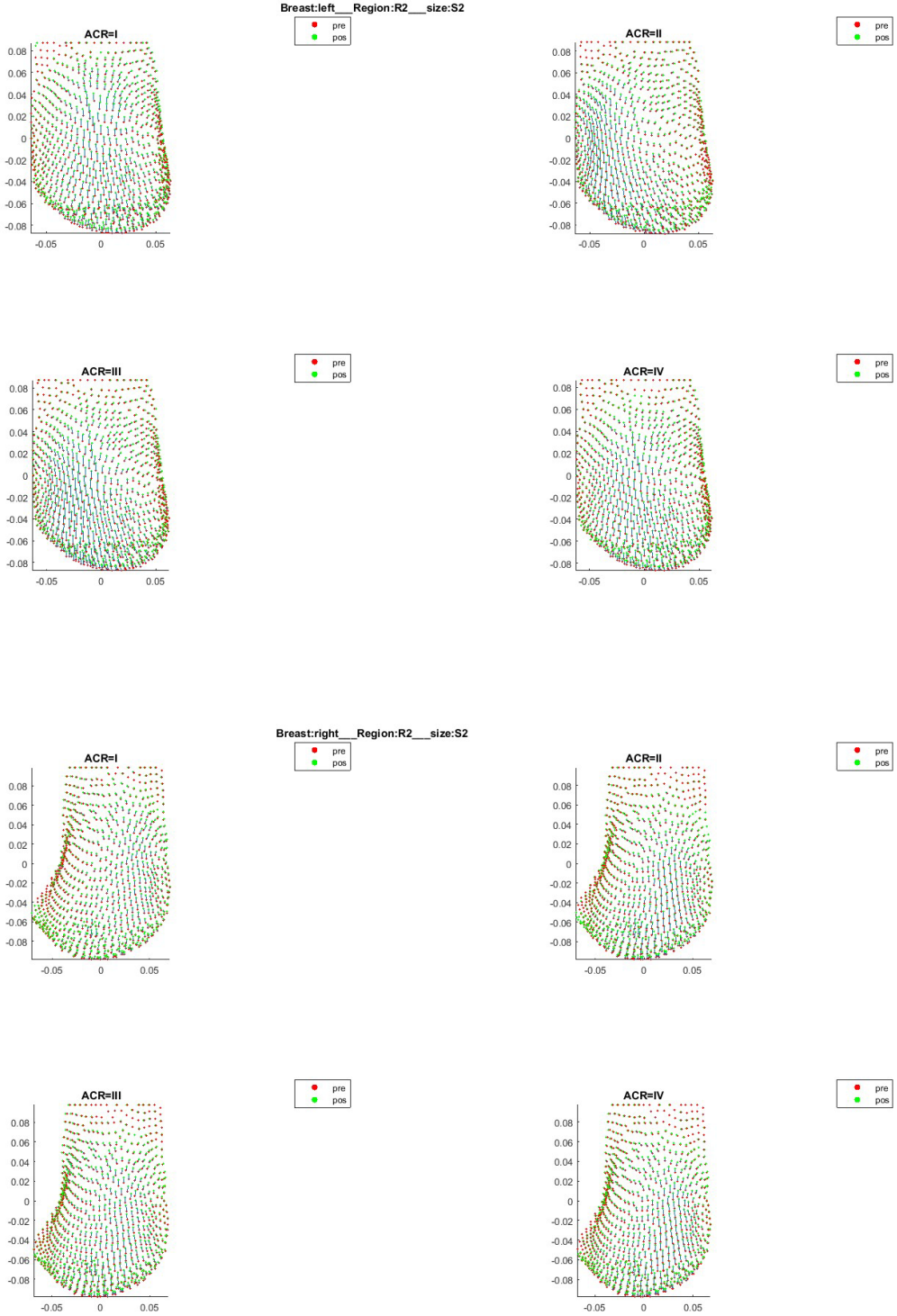
Breast:right___Region:R1___size:S2

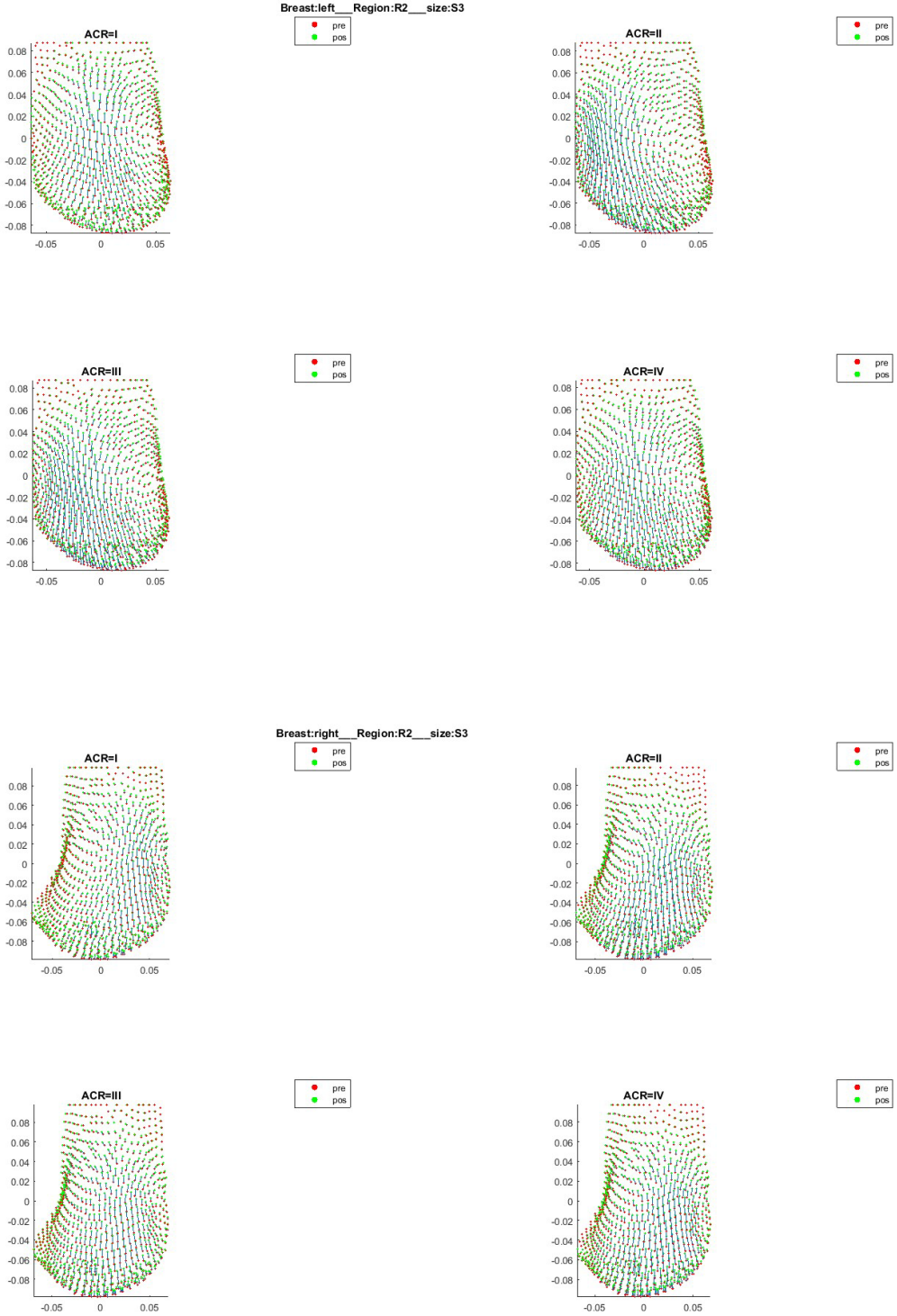


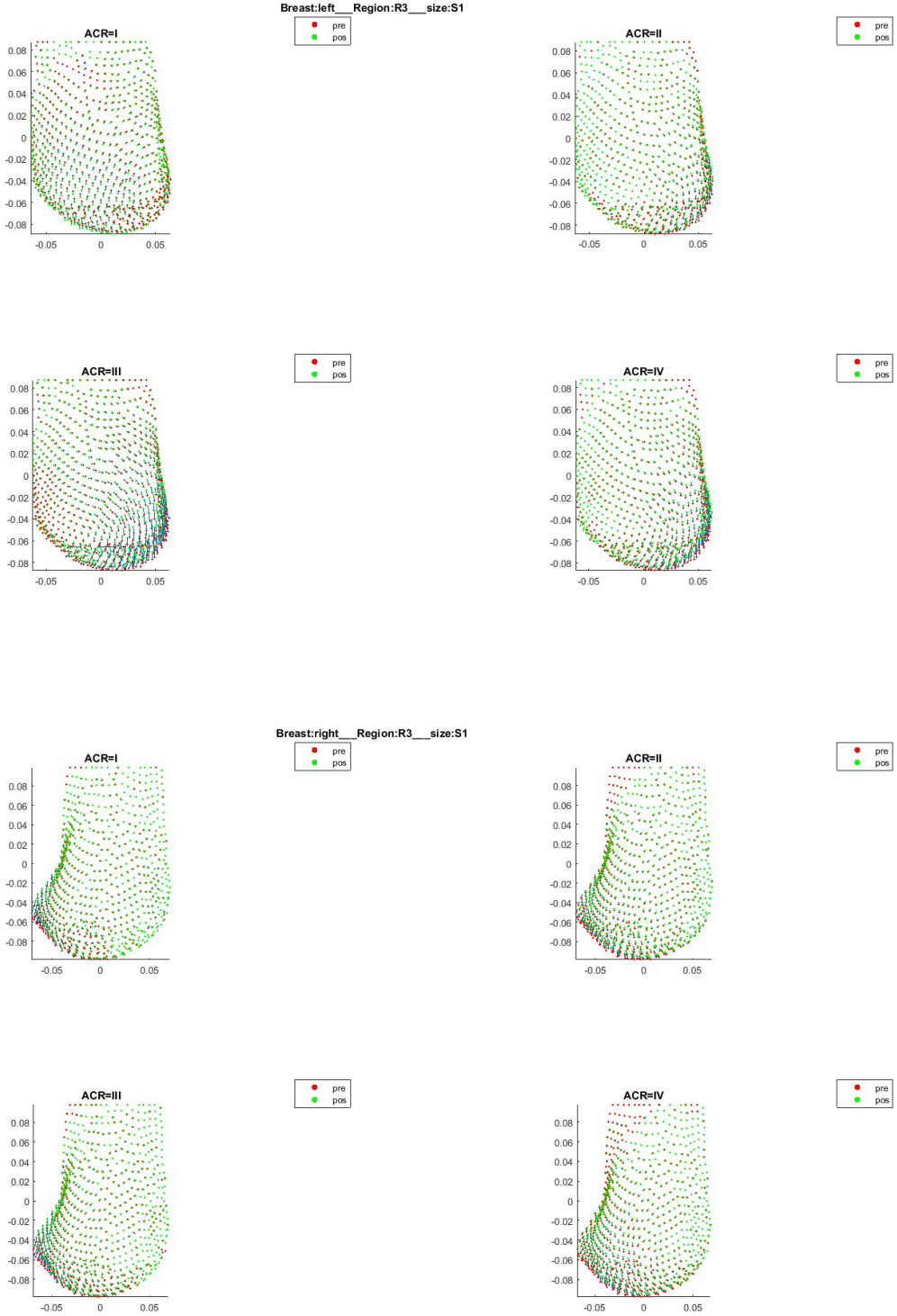


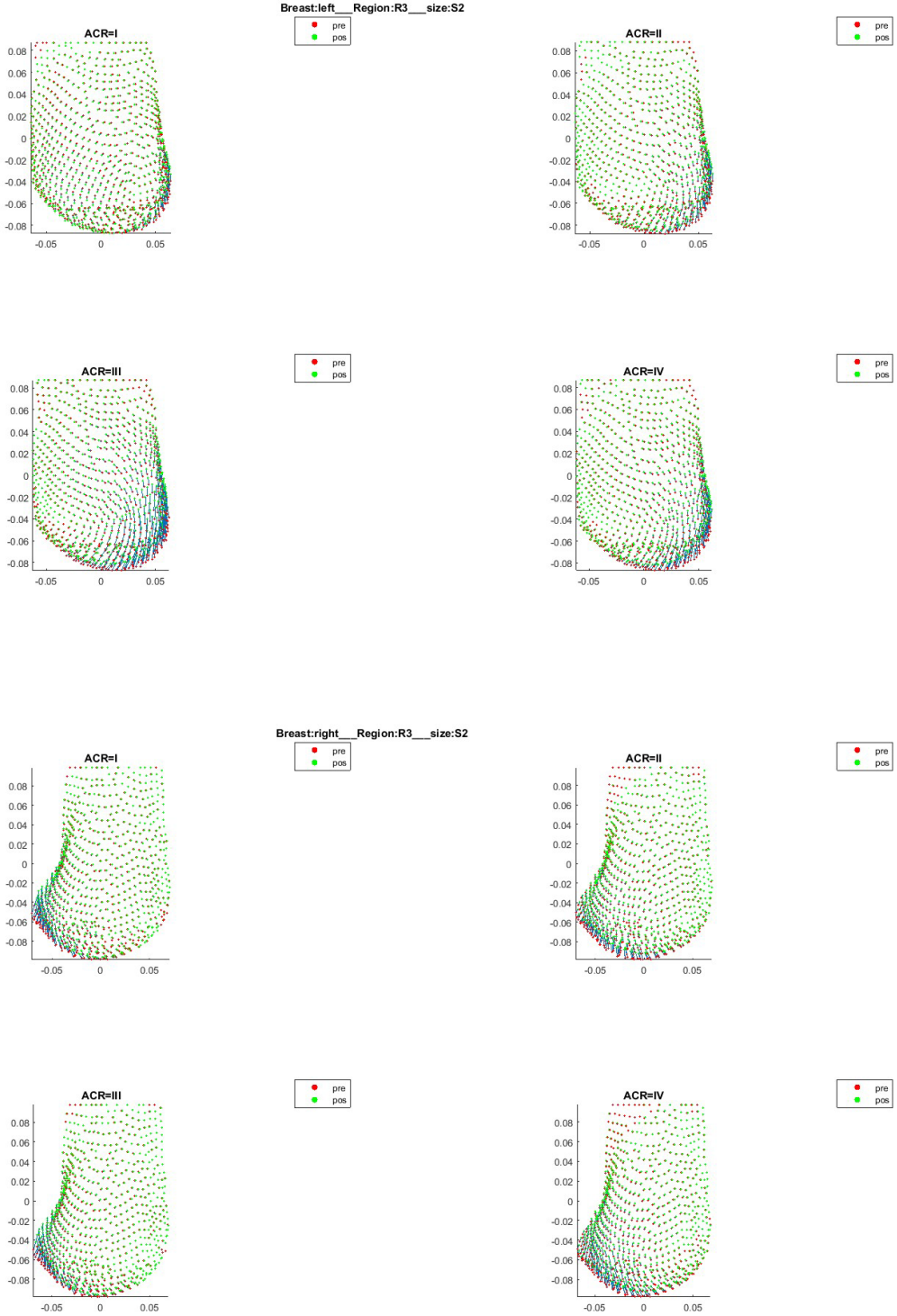

Breast:left___Region:R1___size:S3
ACR=I
ACR=II
ACR=III
ACR=IV
Breast:right___Region:R1___size:S3
ACR=I
ACR=II
ACR=III
ACR=IV
pre
pos

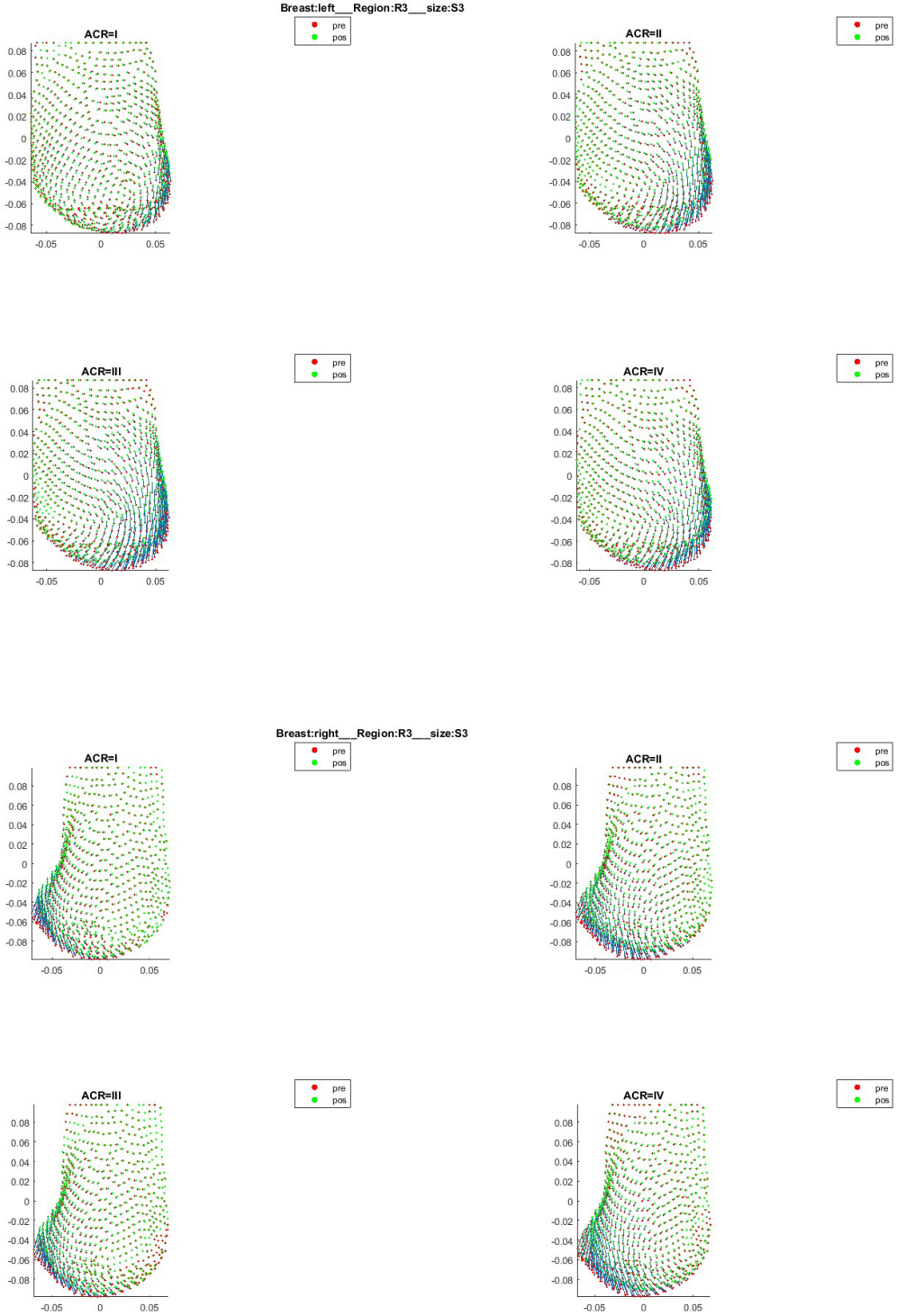

Breast:left___Region:R2___size:S1
ACR=I
ACR=II
ACR=III
ACR=IV
Breast:right___Region:R2___size:S1
ACR=I
ACR=II
ACR=III
ACR=IV
pre
pos


Breast:left___Region:R2___size:S2
ACR=I
ACR=II
ACR=III
ACR=IV
Breast:right___Region:R2___size:S2
ACR=I
ACR=II
ACR=III
ACR=IV
pre
pos

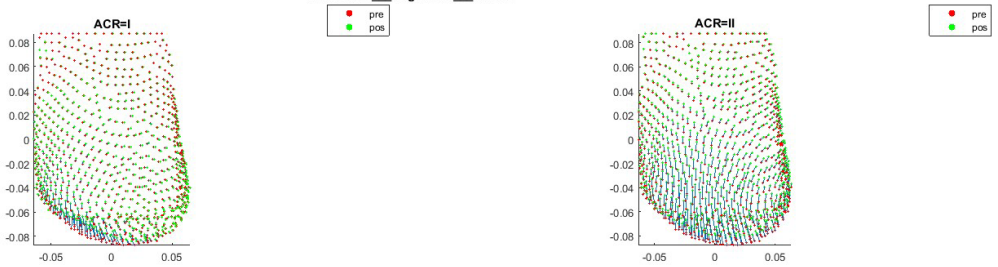

Breast:left___Region:R2___size:S3
ACR=I
ACR=II
ACR=III
ACR=IV
Breast:right___Region:R2___size:S3
ACR=I
ACR=II
ACR=III
ACR=IV
pre
pos

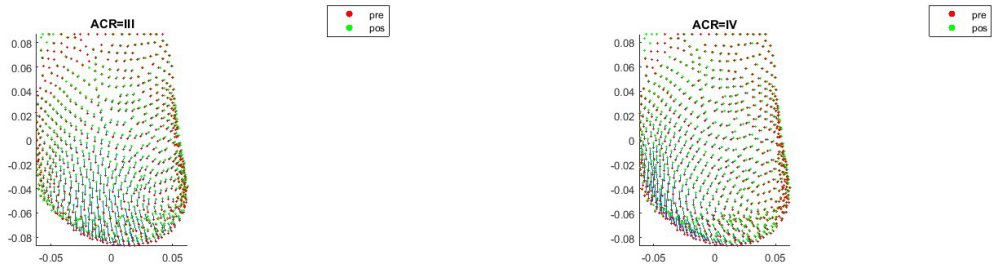

Breast:left___Region:R3___size:S1
ACR=I
ACR=II
ACR=III
ACR=IV
Breast:right___Region:R3___size:S1
ACR=I
ACR=II
ACR=III
ACR=IV
pre
pos
0.08
0.06
0.04
0.02
0
-0.02
-0.04
-0.06
-0.08
-0.05
0
0.05


Breast:left___Region:R3___size:S2
ACR=I
ACR=II
ACR=III
ACR=IV
Breast:right___Region:R3___size:S2
ACR=I
ACR=II
ACR=III
ACR=IV
pre
pos

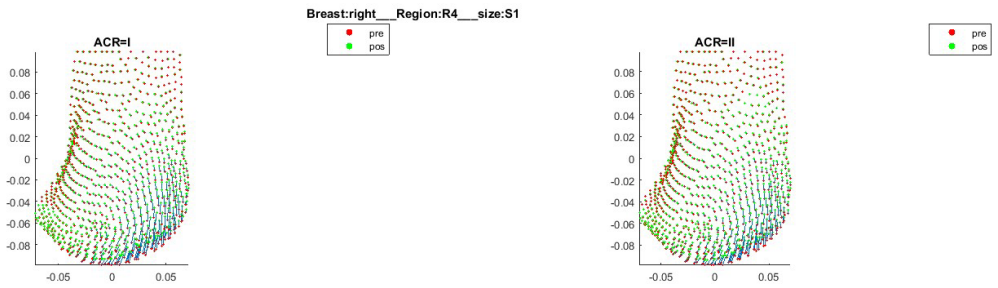

Breast:left___Region:R3___size:S3
ACR=I
ACR=II
ACR=III
ACR=IV
Breast:right___Region:R3___size:S3
ACR=I
ACR=II
ACR=III
ACR=IV
pre
pos
0.08
0.06
0.04
0.02
0
-0.02
-0.04
-0.06
-0.08
-0.05
0
0.05

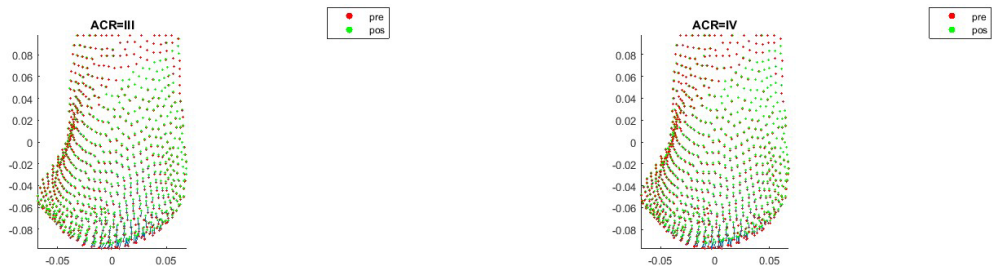
Breast:left___Region:R4___size:S1

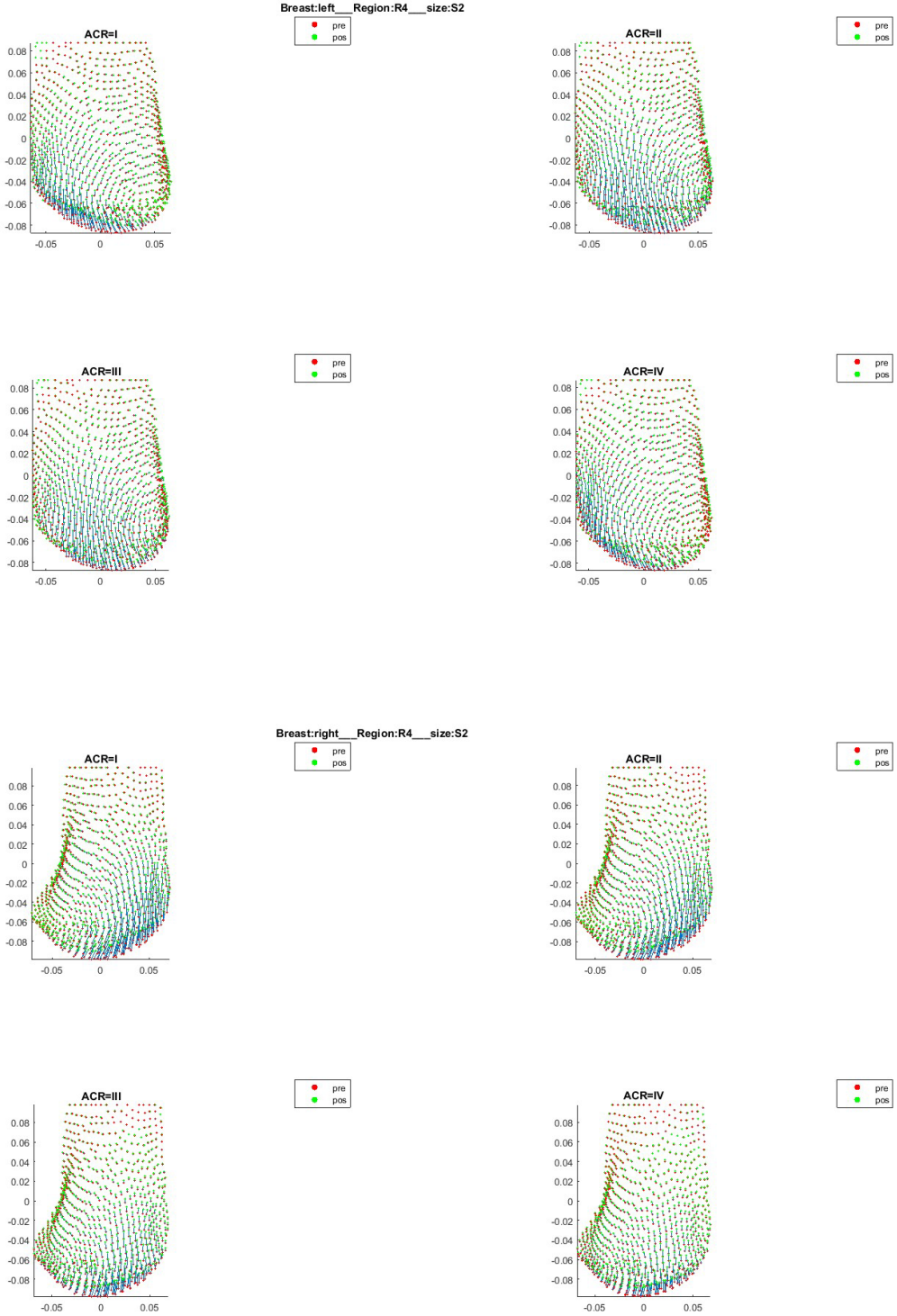


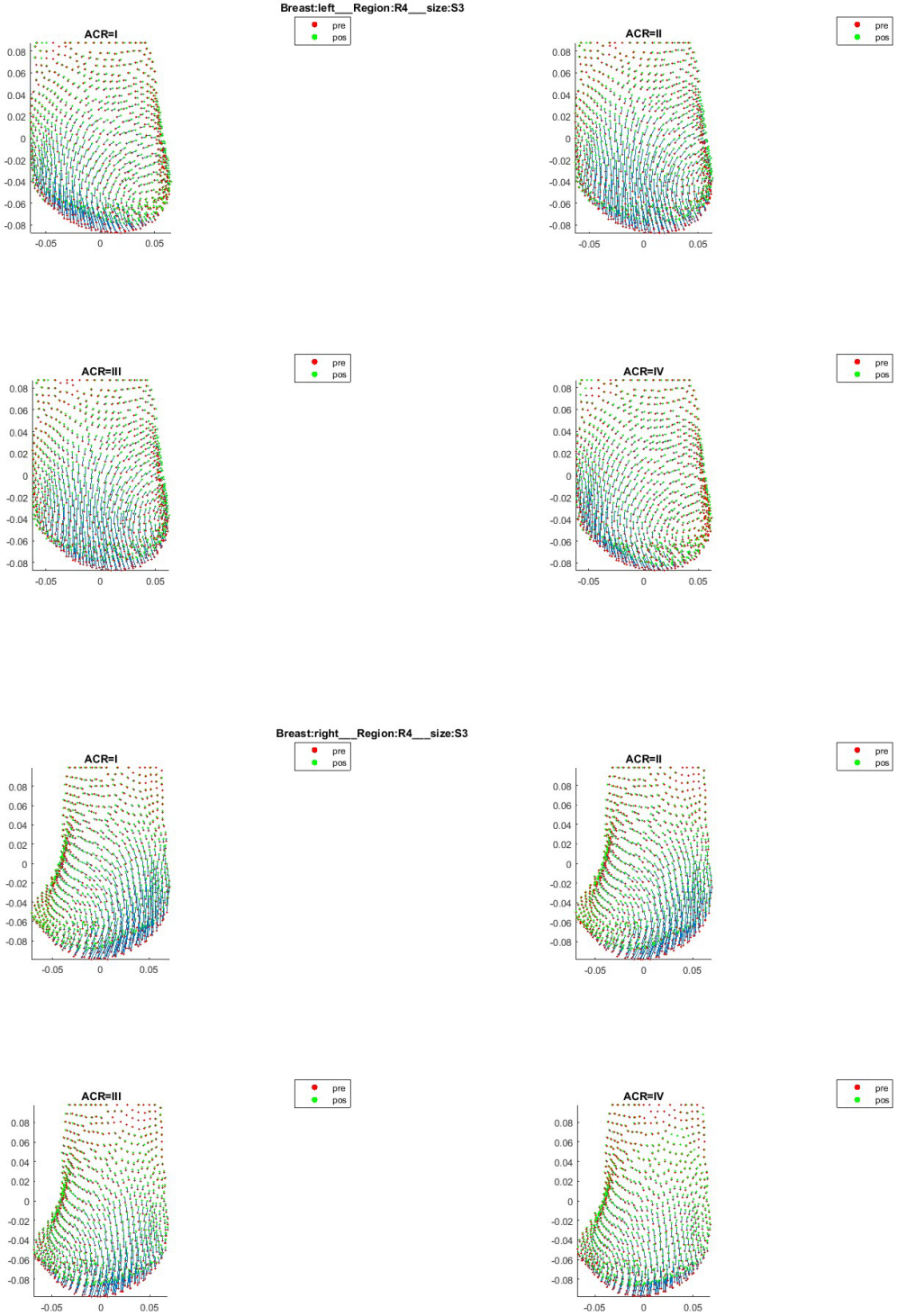


Breast:right___Region:R4___size:S1



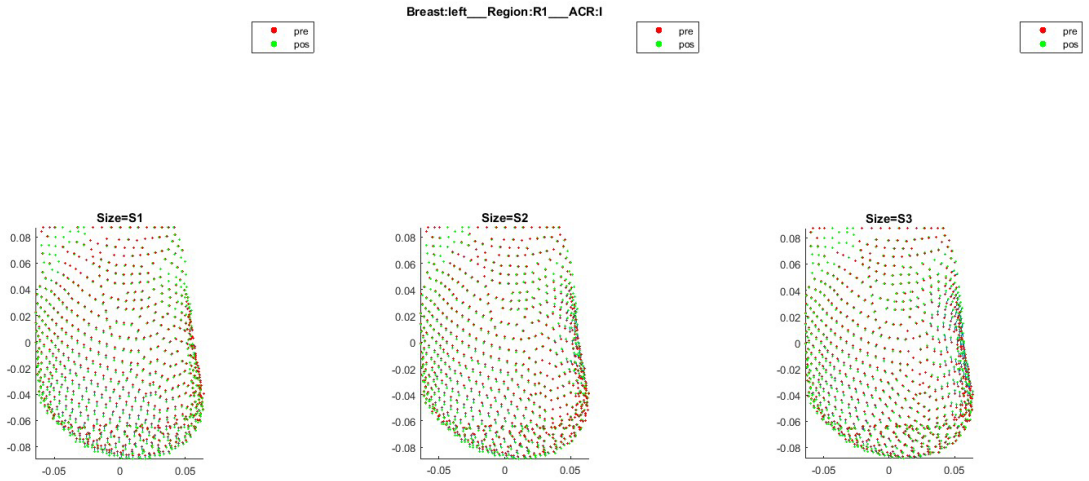


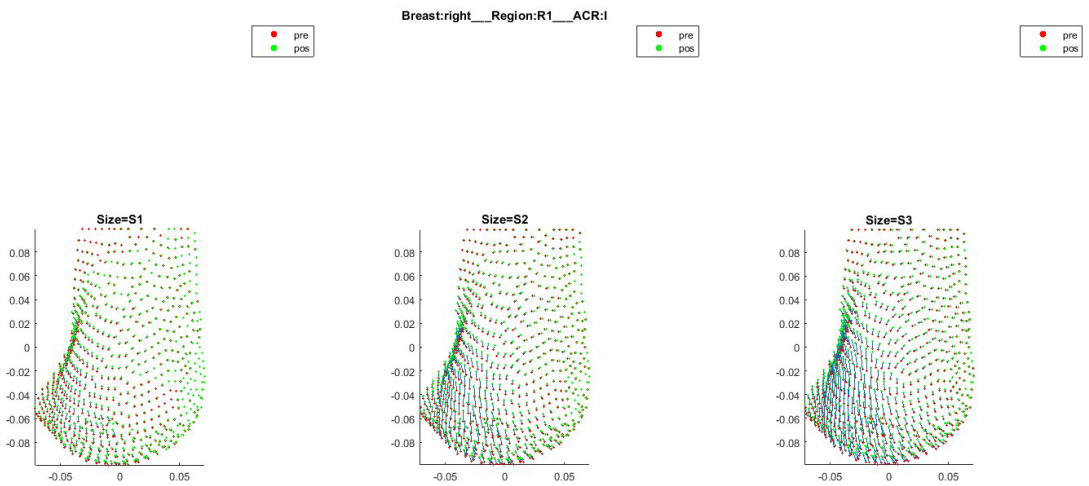

Breast:left___Region:R4___size:S2
ACR=I
ACR=II
ACR=III
ACR=IV
Breast:right___Region:R4___size:S2
ACR=I
ACR=II
ACR=III
ACR=IV
pre
pos

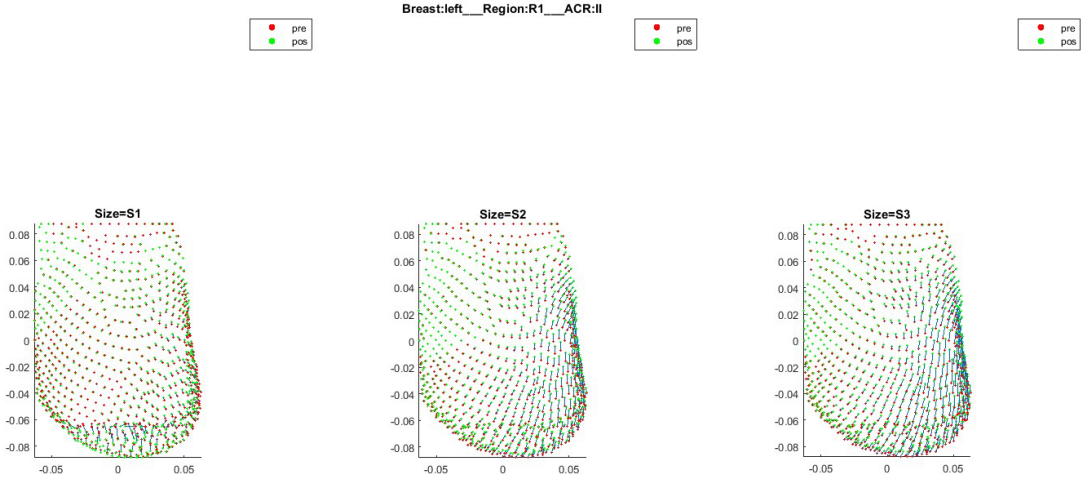

Breast:left___Region:R4___size:S3
ACR=I
ACR=II
ACR=III
ACR=IV
Breast:right___Region:R4___size:S3
ACR=I
ACR=II
ACR=III
ACR=IV
pre
pos

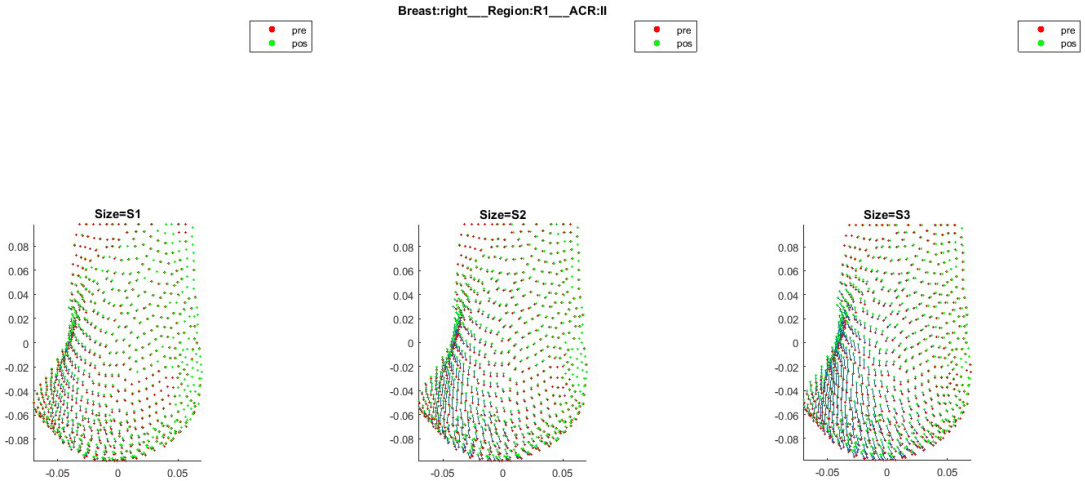
Fixing breast laterality, Region and ACR:

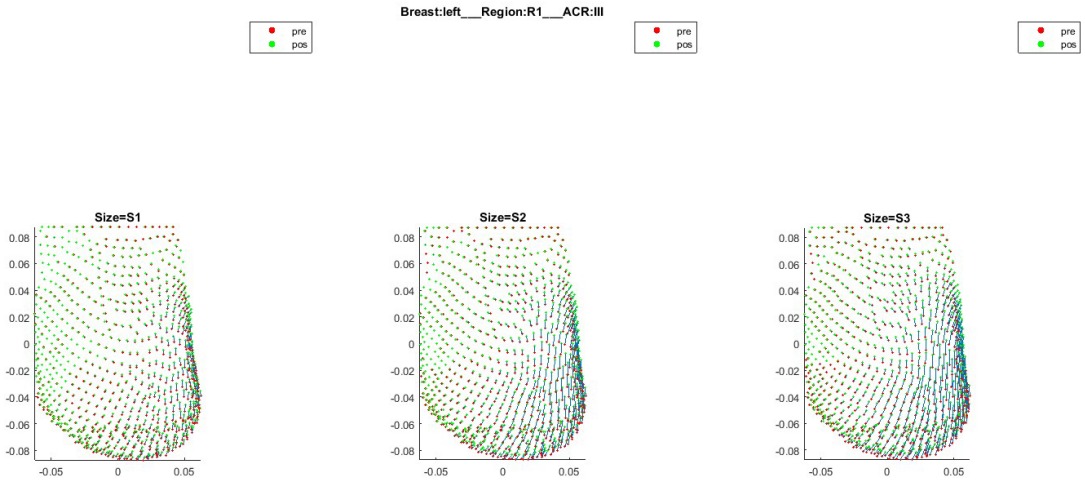
Varying Tumor size

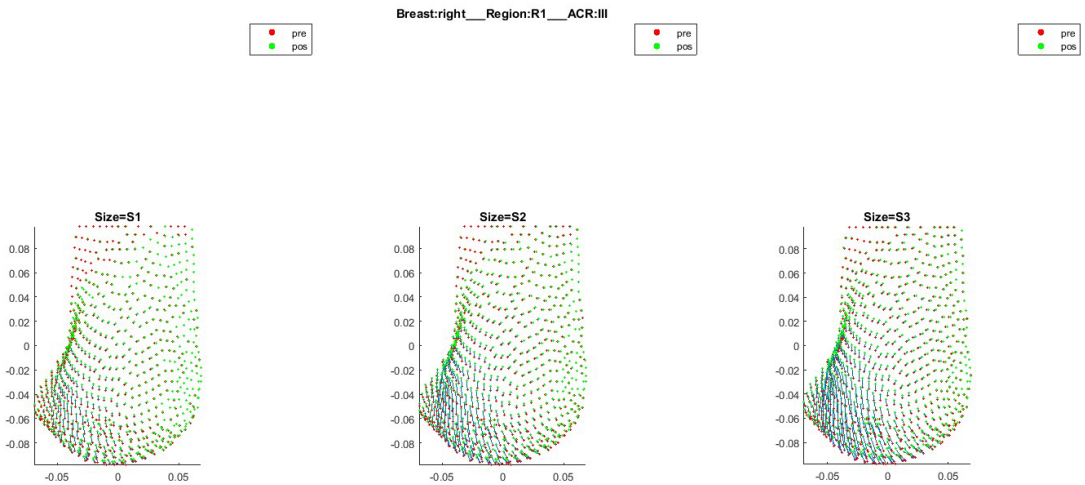


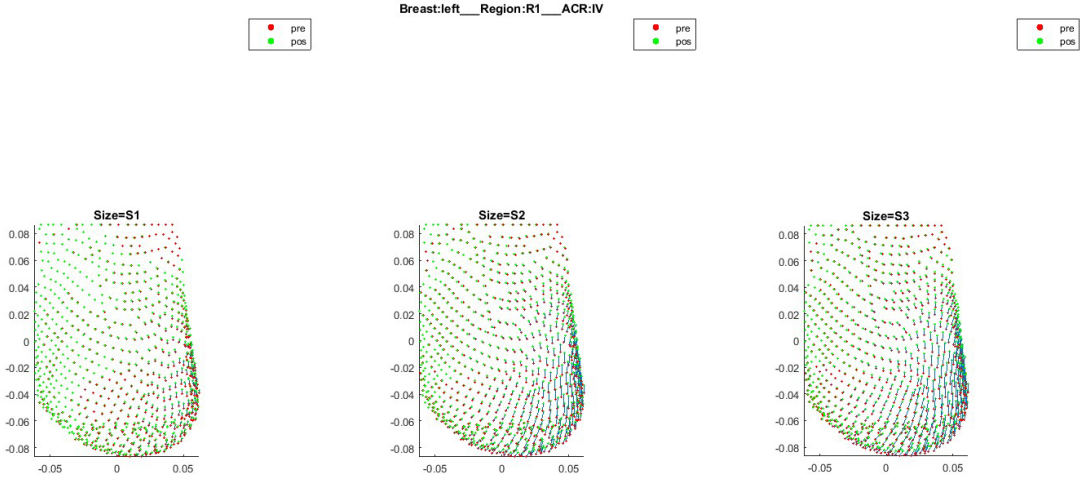


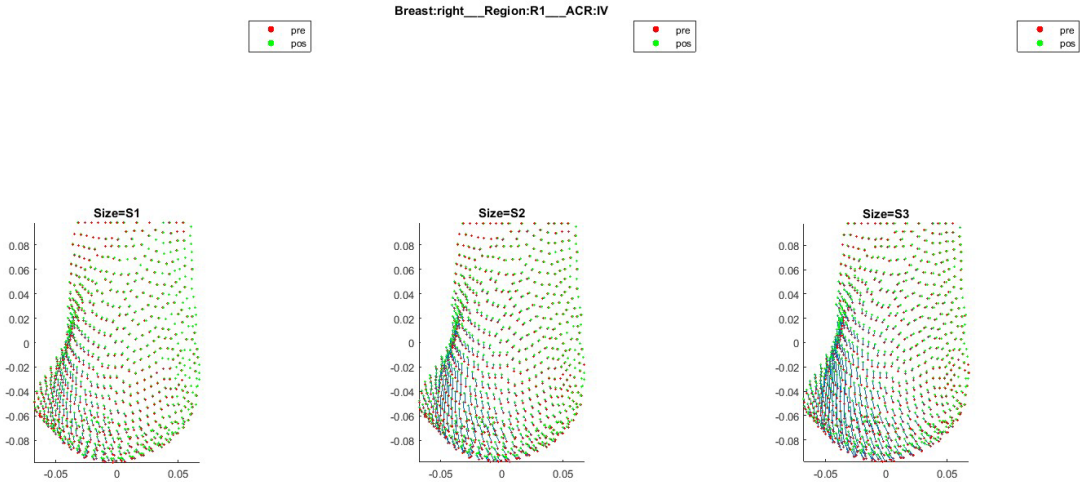

Breast:left___Region:R1___ACR:II
pre
pos
Size=S1
Size=S2
Size=S3

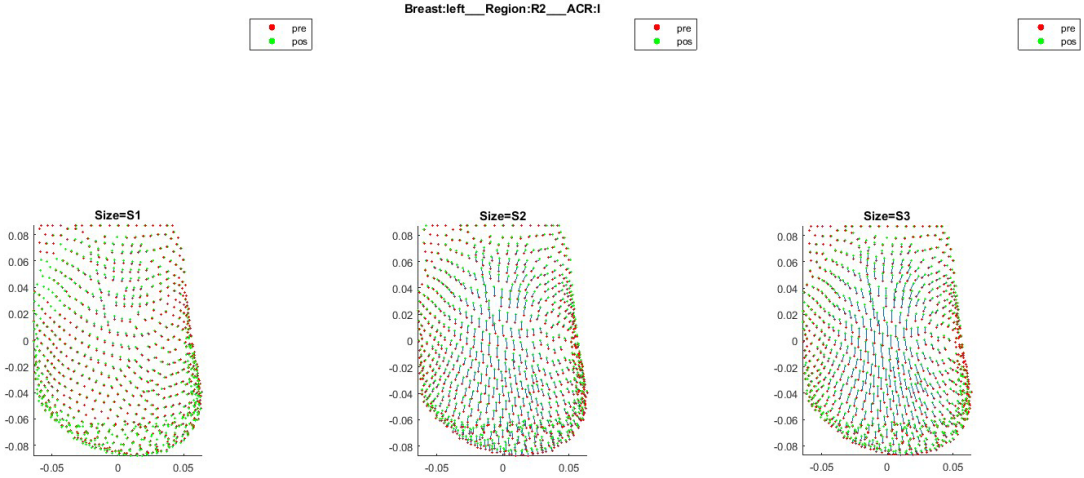

Breast:right___Region:R1___ACR:II
pre
pos
Size=S1
Size=S2
Size=S3

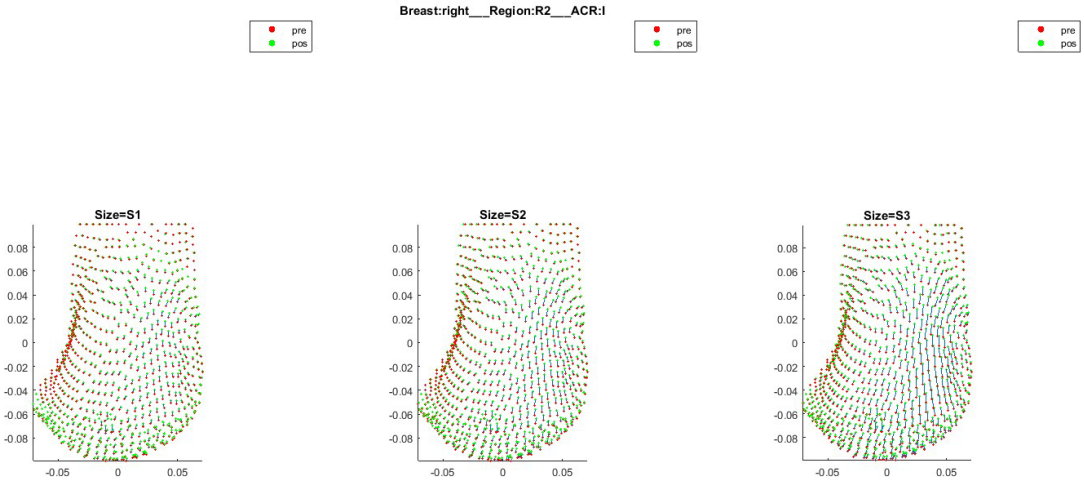

Breast:left___Region:R1___ACR:III
pre
pos
Size=S1
Size=S2
Size=S3

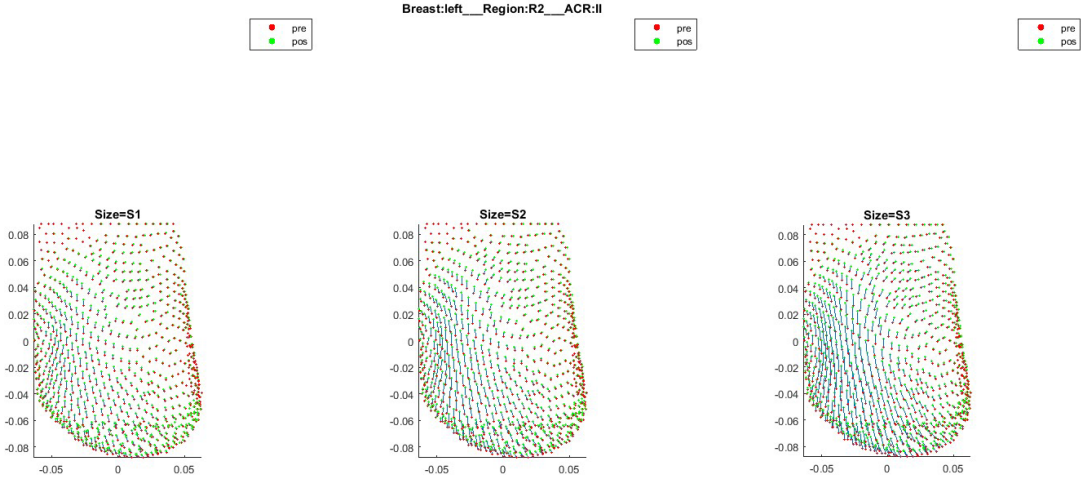

Breast:right___Region:R1___ACR:III
pre
pos
Size=S1
Size=S2
Size=S3

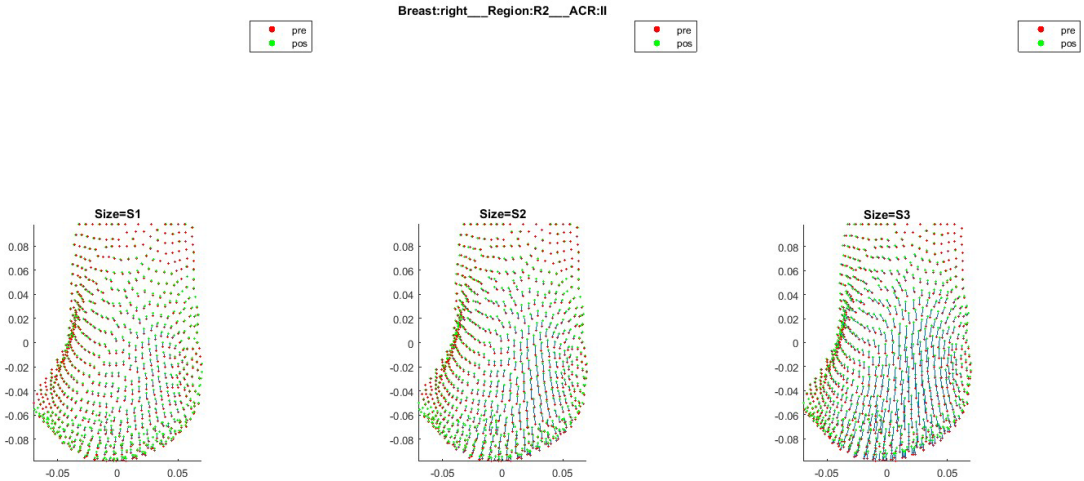

Breast:left___Region:R1___ACR:IV
pre
pos
Size=S1
Size=S2
Size=S3

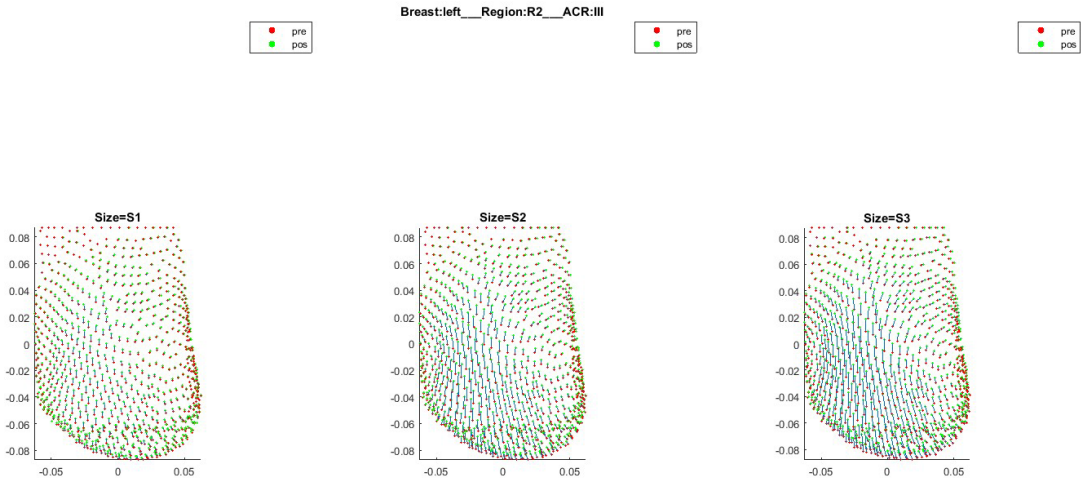

Breast:right___Region:R1___ACR:IV
pre
pos
Size=S1
Size=S2
Size=S3

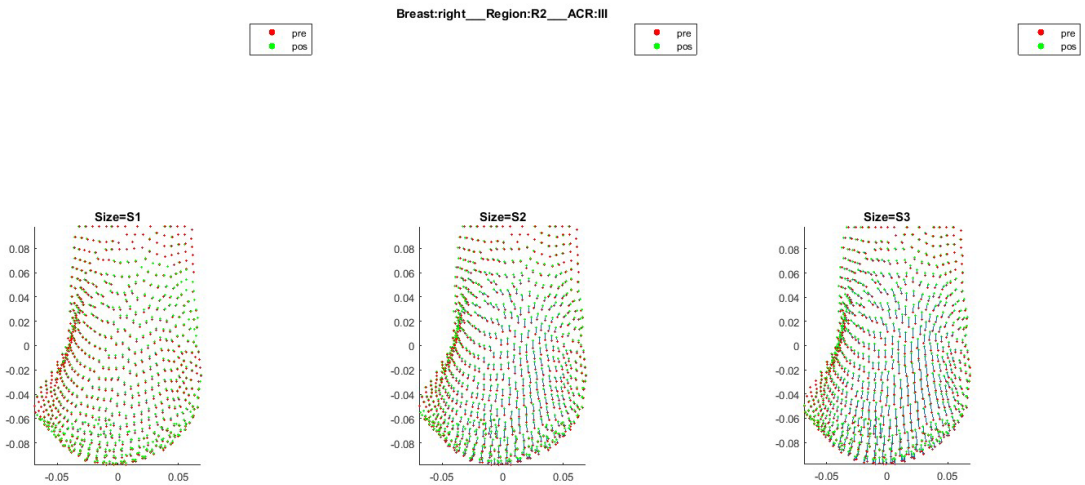

Breast:left___Region:R2___ACR:I
pre
pos
Size=S1
Size=S2
Size=S3

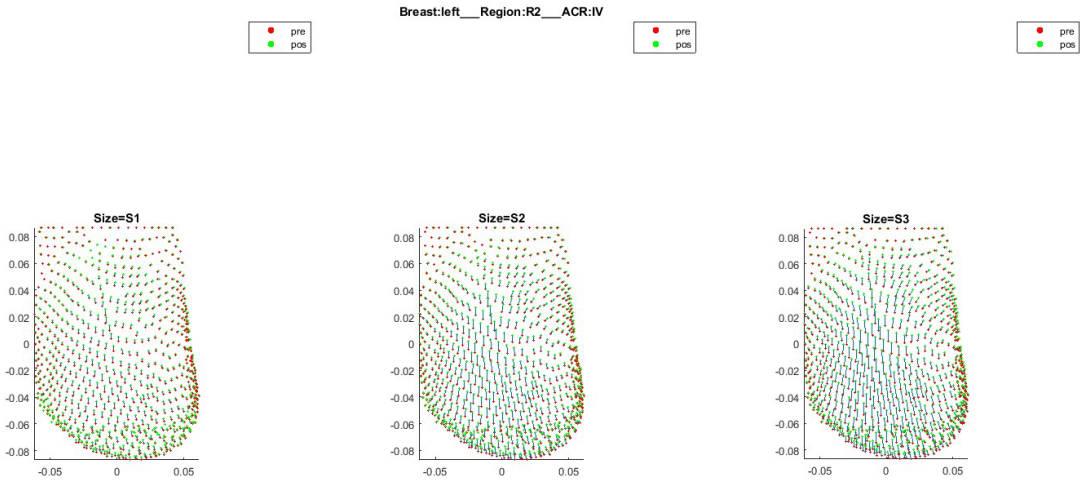

Breast:right___Region:R2___ACR:I
pre
pos
Size=S1
Size=S2
Size=S3

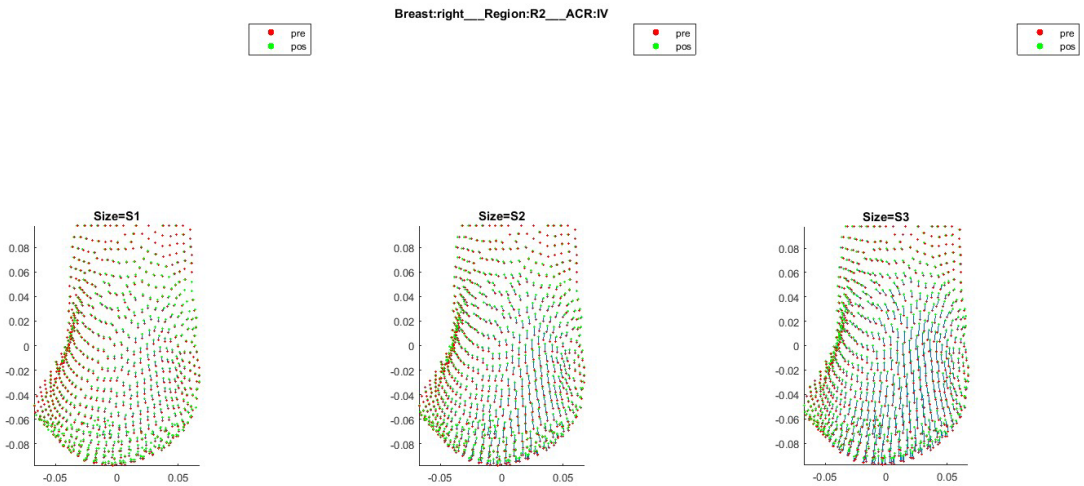

Breast:left___Region:R2___ACR:II
pre
pos
Size=S1
Size=S2
Size=S3

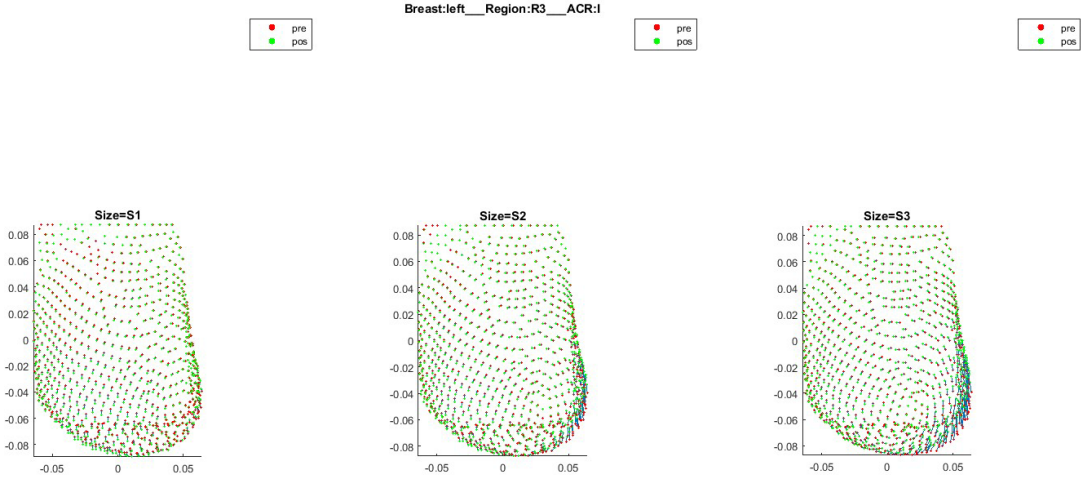

Breast:right___Region:R2___ACR:II
pre
pos
Size=S1
Size=S2
Size=S3

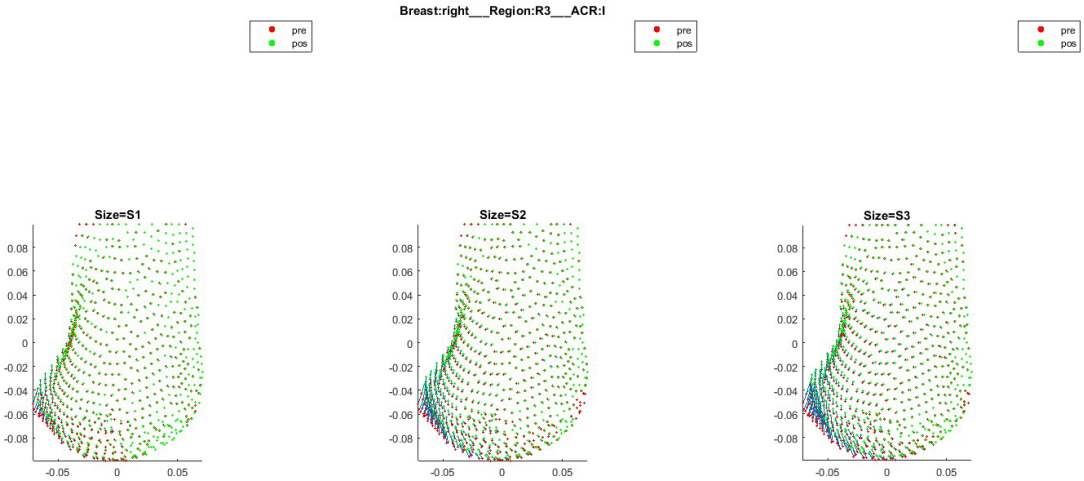

Breast:left___Region:R2___ACR:III
pre
pos
Size=S1
Size=S2
Size=S3

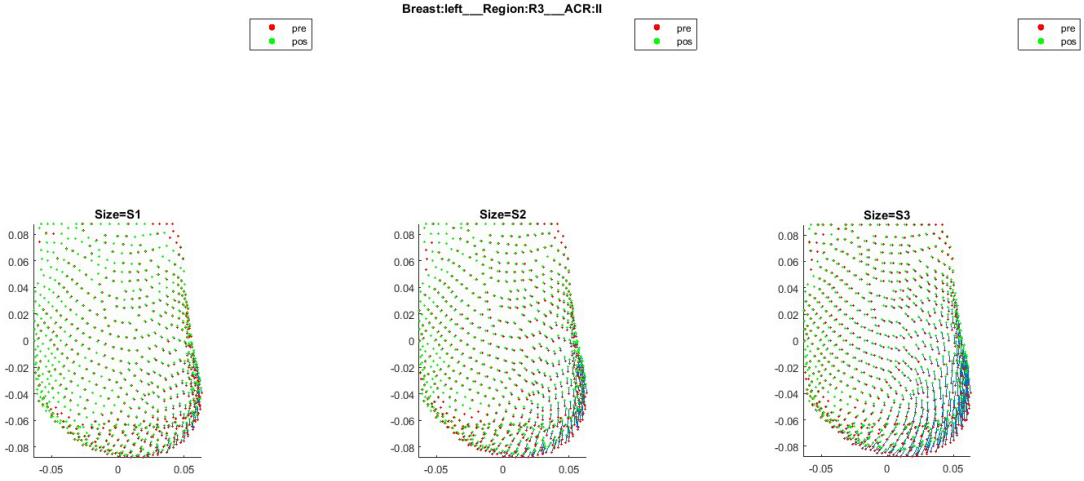

Breast:right___Region:R2___ACR:III
pre
pos
Size=S1
Size=S2
Size=S3

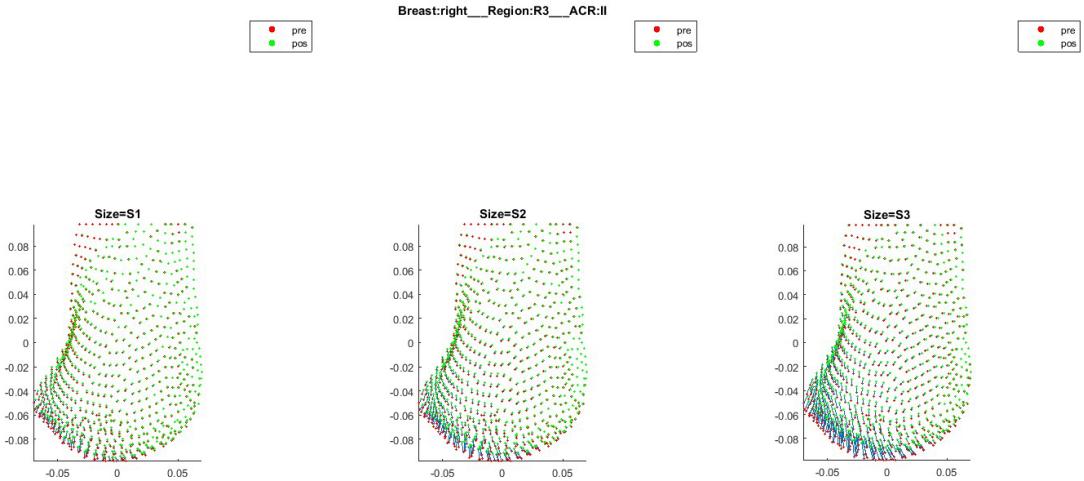

Breast:left___Region:R2___ACR:IV
pre
pos
Size=S1
Size=S2
Size=S3

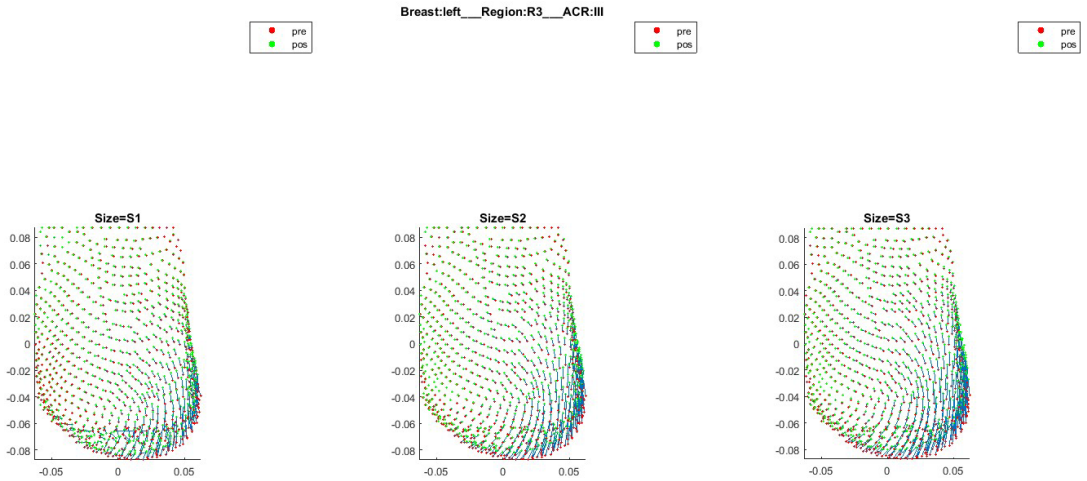

Breast:right___Region:R2___ACR:IV
pre
pos
Size=S1
Size=S2
Size=S3

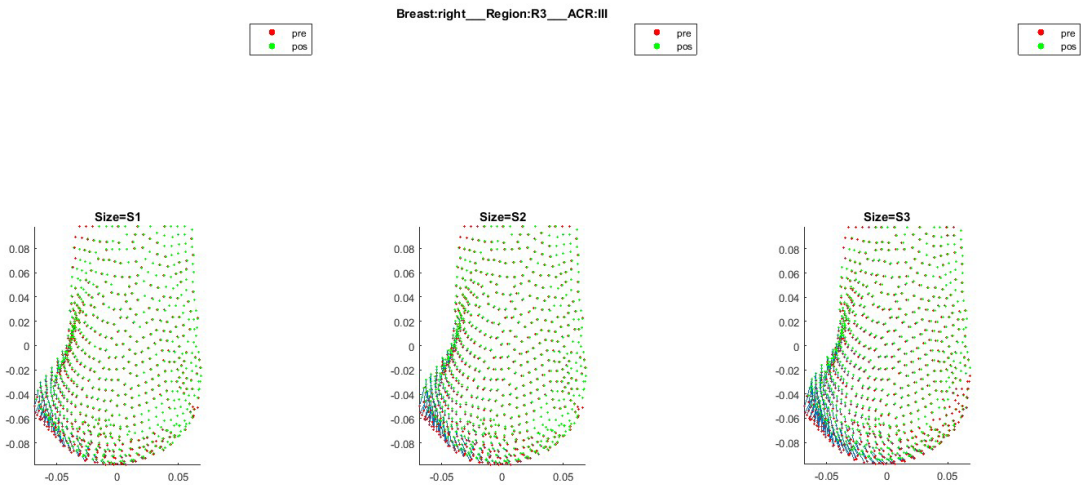

Breast:left___Region:R3___ACR:I
pre
pos
Size=S1
Size=S2
Size=S3

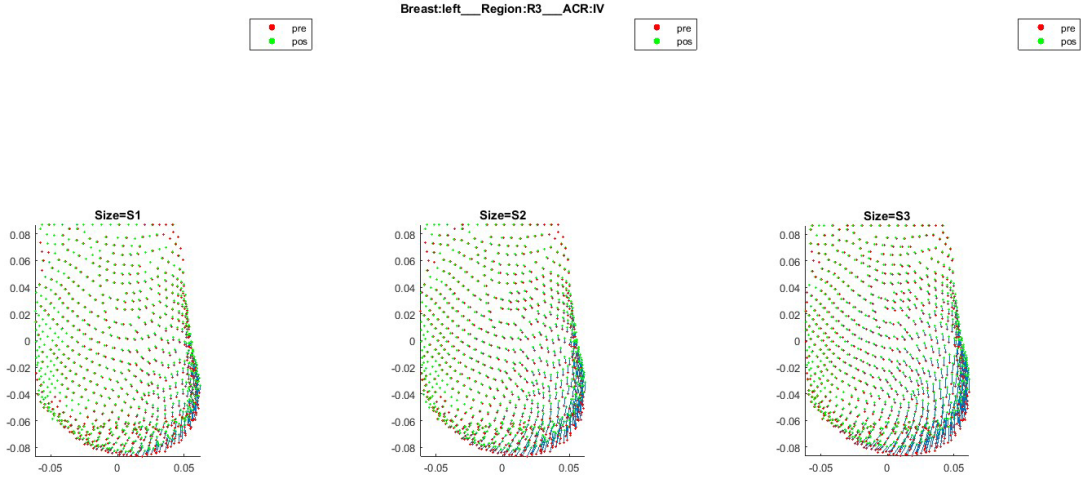

Breast:right___Region:R3___ACR:I
pre
pos
Size=S1
Size=S2
Size=S3

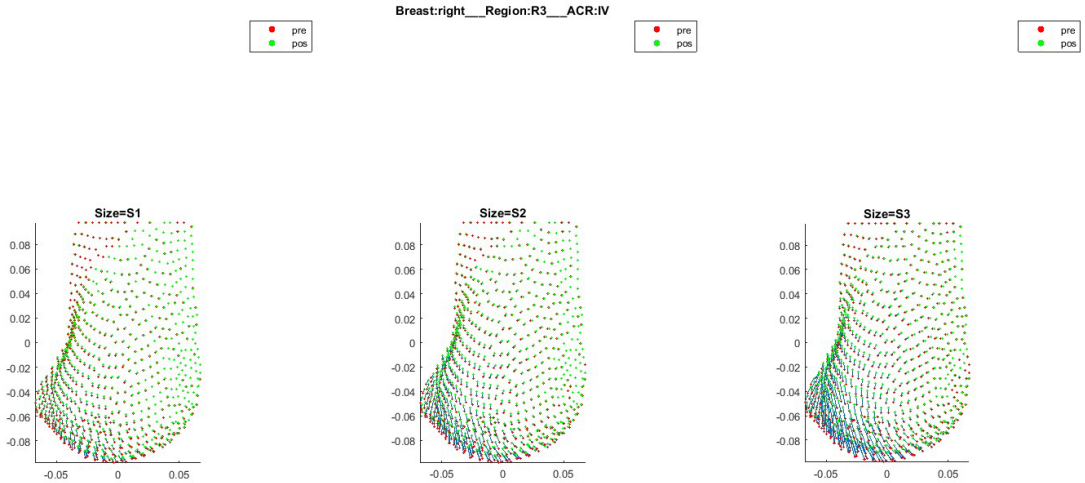

Breast:left___Region:R3___ACR:II
pre
pos
Size=S1
Size=S2
Size=S3

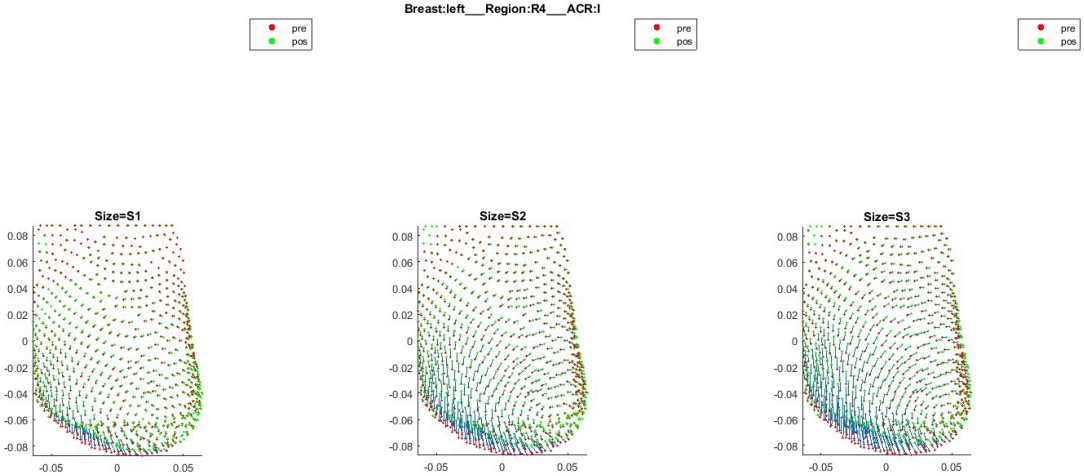

Breast:right___Region:R3___ACR:II
pre
pos
Size=S1
Size=S2
Size=S3

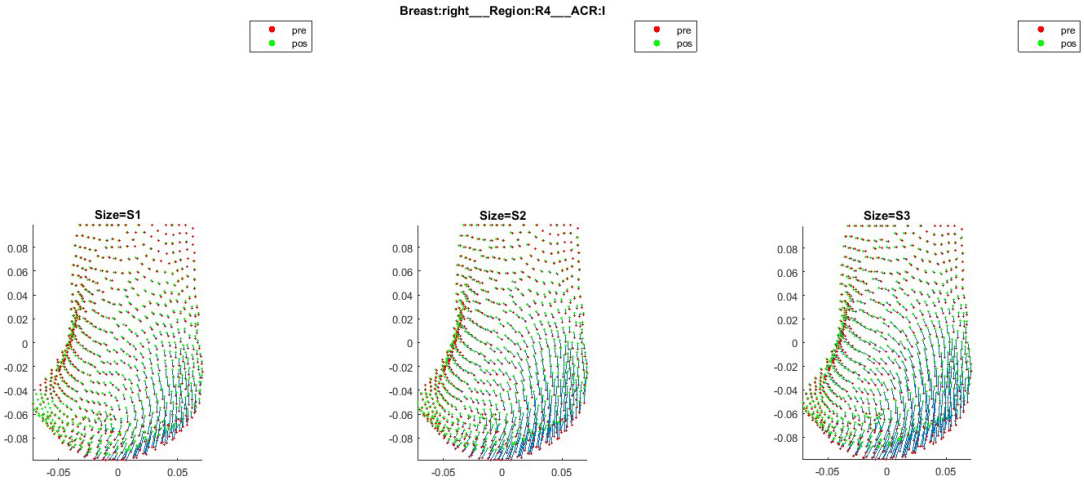

Breast:left___Region:R3___ACR:III
pre
pos
Size=S1
Size=S2
Size=S3

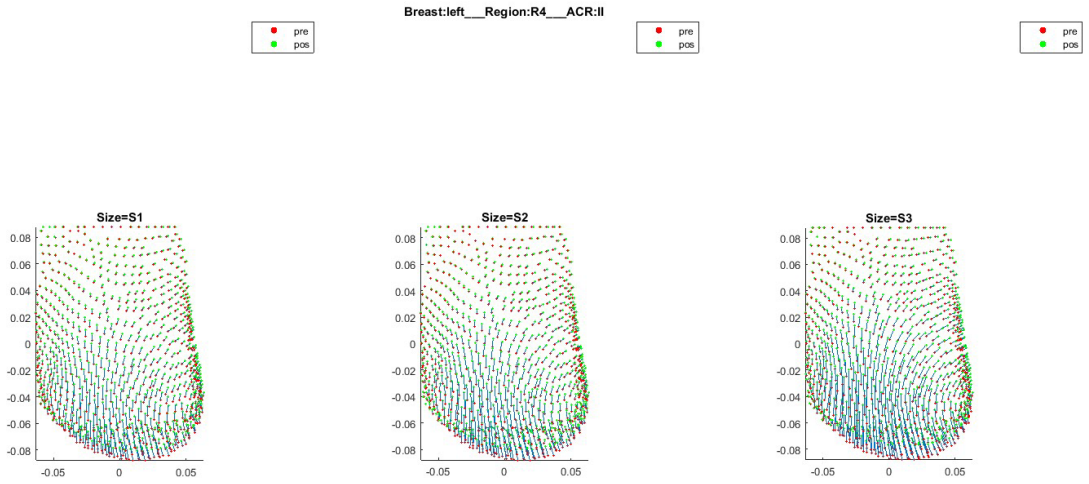

Breast:right___Region:R3___ACR:III
pre
pos
Size=S1
Size=S2
Size=S3

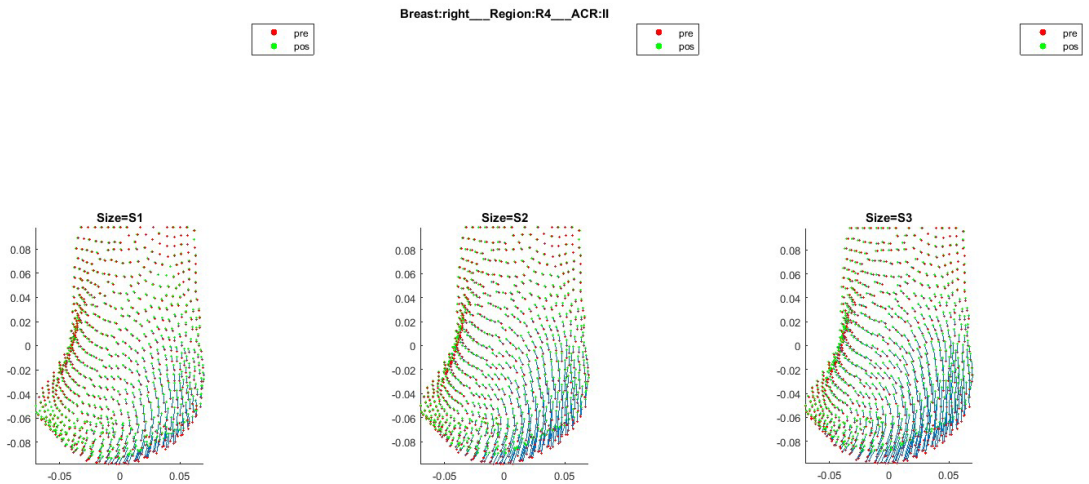

Breast:left___Region:R3___ACR:IV
pre
pos
Size=S1
Size=S2
Size=S3

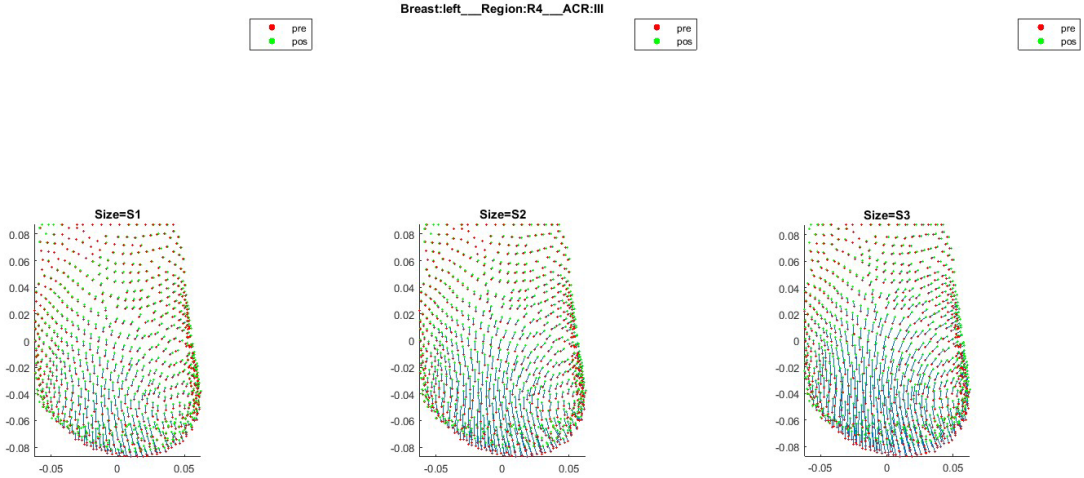

Breast:right___Region:R3___ACR:IV
pre
pos
Size=S1
Size=S2
Size=S3

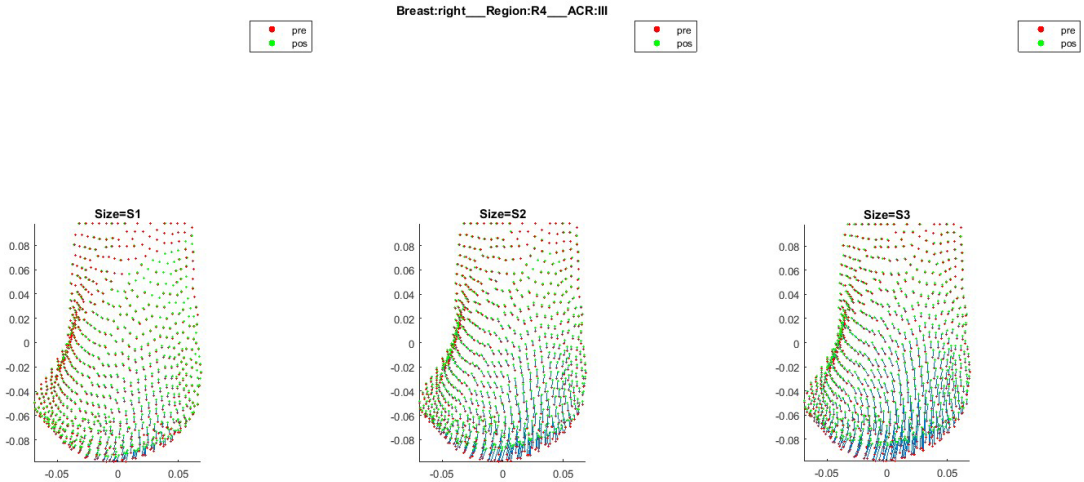

Breast:left___Region:R4___ACR:I
pre
pos
Size=S1
Size=S2
Size=S3

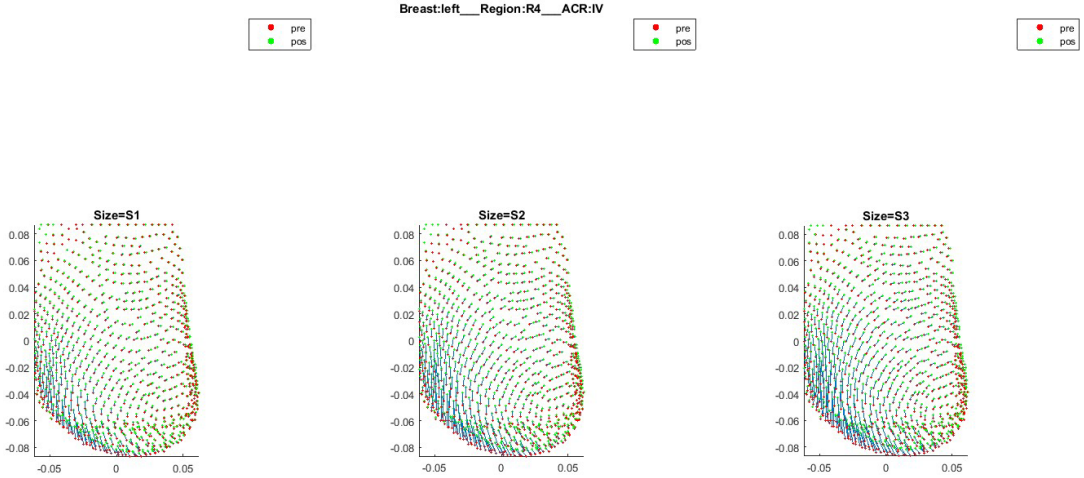

Breast:right___Region:R4___ACR:I
pre
pos
Size=S1
Size=S2
Size=S3

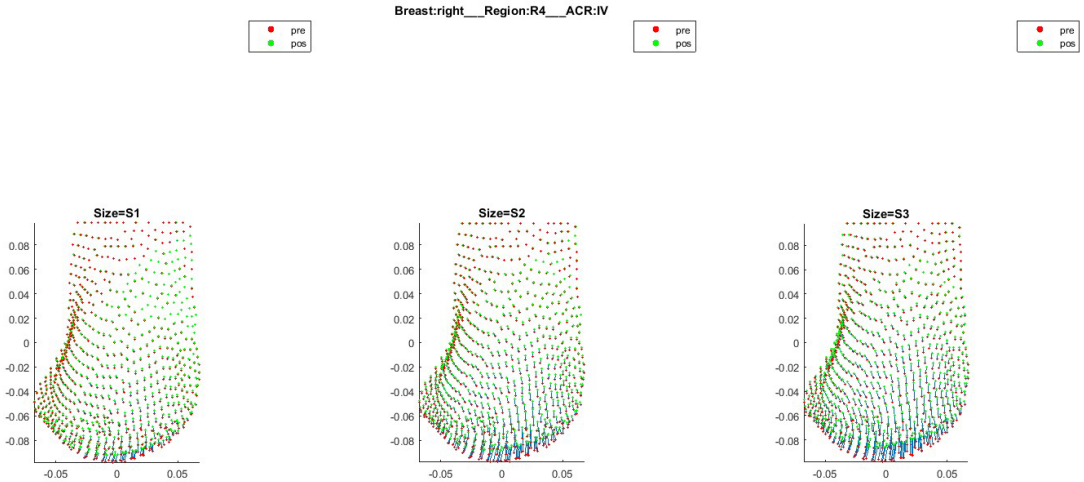

Breast:left___Region:R4___ACR:II
pre
pos
Size=S1
Size=S2
Size=S3


Breast:right___Region:R4___ACR:II
pre
pos
Size=S1
Size=S2
Size=S3


Breast:left___Region:R4___ACR:III
pre
pos
Size=S1
Size=S2
Size=S3


Breast:right___Region:R4___ACR:III
pre
pos
Size=S1
Size=S2
Size=S3


Breast:left___Region:R4___ACR:IV
pre
pos
Size=S1
Size=S2
Size=S3


Breast:right___Region:R4___ACR:IV
pre
pos
Size=S1
Size=S2
Size=S3